

Accession	Length (bp)	GC (%)	Gene
1	494	100.0	AX472612
2	494	100.0	147709 <i>Bacillus subtilis</i> 29915
3	494	100.0	AX100815
4	485	99.2	AX100815
5	258	52.2	AX100850
6	257	52.0	AX572874
7	257	52.0	AX573226
8	257	52.0	AX100849
9	257	52.0	AX100832
10	244	49.4	AX100779
11	229	46.4	AXE017333
12	229	46.4	CP000002
13	216.2	43.8	AX572872
14	216.2	43.8	AX573224
15	203	41.1	BACJ0269
16	161.8	32.8	AXE017269
17	160.2	32.4	AXE017355
18	160.2	32.4	AXE017002
19	158.6	32.1	AXE017225

C	20	158.6	32.1	110000	1	AE017334_14	Continuation (15
C	21	158.6	32.1	110000	1	CP000001_14	Continuation (15
C	22	158.6	32.1	110000	1	CP000001_15	Continuation (16
C	23	158.6	32.1	290525	1	AE017029	AE017029 Bacillus
C	24	151.2	30.6	834	6	AX432591	AX432591 Sequence
C	25	137	27.7	1613	6	AX416370	AX416270 Sequence
C	26	137	27.7	313450	1	AL596170	AL596170 Listeria
C	27	137	27.7	319630	6	AX413016	AX413016 Sequence
C	28	137	27.7	349980	6	AX417046	AX417046 Sequence
C	29	133.8	27.1	291954	1	AE017328	AE017328 Listeria
C	30	133.8	27.1	347050	1	AL591981	AL591981 Listeria
C	31	133.8	27.1	349980	6	AX416170	AX416170 Sequence
C	32	117.2	23.7	1880	1	AF270386	AF270386 Staphyloc
C	33	117.2	23.7	1880	6	AR486340	AR486340 Sequence
C	34	117.2	23.7	1880	6	AX145704	AX145704 Sequence
C	35	117.2	23.7	300275	1	AE016751	AE016751 Staphyloc
C	36	114.4	23.2	300550	1	AP001512	AP001512 Bacillus
C	37	103.2	20.9	792	6	AR484290	AR484290 Sequence
C	38	103.2	20.9	792	6	AX142655	AX142655 Sequence
C	39	99.6	20.2	4549	6	AR535612	AR535612 Sequence
C	40	99.6	20.2	110000	1	BX571856_27	Continuation (28
C	41	99.6	20.2	110000	1	BX571857_26	Continuation (27
C	42	99.6	20.2	110000	1	AP003137	AP003137 Staphyloc
C	43	99.6	20.2	297850	1	AP004830	AP004830 Staphyloc
C	44	99.6	20.2	301250	1	AP003365	AP003365 Staphyloc
C	45	98	19.8	342350	1	AP003365	AP003365 Staphyloc

ALIGNMENTS

RESULT 1	AX472612	494 bp	DNA	linear	PAT 09-AUG-2002
LOCUS	AX472612				
DEFINITION	Sequence 1 from Patent WO0214559.				
ACCESSION	AX472612				
VERSION	AX472612.1				
KEYWORDS	GI:22207504				
SOURCE					
ORGANISM	Bacillus subtilis				
	Bacillus subtilis				
	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.				
REFERENCE	1				
AUTHORS	Murphy, C.				
TITLE	High throughput screen for inhibitors of the folate biosynthetic pathway in bacteria				
JOURNAL	Patent: WO 0214559-A 1 21-FEB-2002;				
	Millennium Pharmaceuticals, Inc. (US)				
FEATURES	Location/Qualifiers				
source	1..494				
	/organism="Bacillus subtilis"				
	/mol_type="unassigned dna"				
	/db_xref="taxon:1423"				

ORIGIN

Query Match	100.0%	Score 494	DB 6	Length 494
Similarity	100.0%	Pred. No. 4,3e-123		
Best Local				
Matches	494	Conservative 0	Mismatches 0	Indels 0
Gaps	0			
QY	1	GCTATGTTGTTGGTAAAGCCCGTTGATTTTGGTATACCTTCATTGGGACGATATGCGCTG	60	
DB	1	GCTATGTTGTTGGTAAAGCCCGTTGATTTTGGTATACCTTCATTGGGACGATATGCGCTG	60	
QY	61	CGAAGTGAAGCTATTAATTAATTAATGATGACACTTGCACAGCTGCTGTGATCCAAAAA	120	
DB	61	CGAAGTGAAGCTATTAATTAATTAATGATGACACTTGCACAGCTGCTGTGATCCAAAAA	120	
QY	121	GGACTGGACAGAGGATGAACTCGCCGAACTTTGAAAGTGAAGATTCCTTCTGCTTG	180	
DB	121	GGACTGGACAGAGGATGAACTCGCCGAACTTTGAAAGTGAAGATTCCTTCTGCTTG	180	
QY	181	TAAAGGAGTTTGGCTTGGCTGCAAGAAAAACGGCAATCATCTCCCTTAACATGAGG	240	
DB	181	TAAAGGAGTTTGGCTTGGCTGCAAGAAAAACGGCAATCATCTCCCTTAACATGAGG	240	

/standard_name="cytochrome b6"
/note="42.3% of identical aminoacids with Chlorella
proteohocoides chloroplast cytochrome b6"
/citation={2}
/codon_start=1
/evidence=experimental
/transl_table=1
/product="cytochrome b"
/protein_id="AAB38436.1"
/db_xref="GI:1146228"
/translation="MLNKIYDWDERLDITPMWRDIDADHEVPEHVNPAHFSAFYCF
GGLRPPPLTVQIVSGMFLTWYYVDPIDKNAESVYLIONEVAFGOIVRGHMWASLVI
VMFHLIRLPFPQGAYKKPRRLNIYGVLFFPMMIGLGFTGYLLIPMDAKALPKTKXL
QIAATRPILGVQTALLAGHPDIVGAQTLTFRFAIHVFILPALFGLMANAFIMIKQ
GISGPL"

gene

CDS

4088..4855
/gene="qcrC"
4088..4855
/gene="qcrC"
/note="29.1% of identical aminoacids with the Spinacia
oleracea chloroplast cytochrome b"
/citation={2}
/codon_start=1
/evidence=experimental
/transl_table=1
/product="cytochrome b"
/protein_id="AAB38437.1"
/db_xref="GI:1146229"
/translation="MHKGKGMKFVGDSRIPIAEKKNIKPDXYSEYPGTETAPMFNFIKL
EMMGAVNLIFGLVTITYTHOPPLERMPDPDTDTYTLPDMYPLELYOLLAKEVAASF
TIVGAMIMPGLAFGLLIAPPDLDRGETERRRKPFVANVMILLISAAPVPTNQSVATH
DVAAKEEGCKTTKENDDIPTNAEGRKVFEQCISCHGDNTQQGAAPSIVDSGLKDE
IKKIAXVEBKQMCPAGVFVKXNDKOIEBLAKFISETTAK"

terminator

4898..4922
/standard_name="transcription terminator"
/note="putative"
/citation={2}

gene

CDS

5002..5559
/gene="ypjA"
5002..5559
/gene="ypjA"
/function="hypothetical"
/note="putative"
/citation={2}
/codon_start=1
/transl_table=1
/protein_id="AAB38438.1"
/db_xref="GI:1146230"
/translation="MLIIIVLAINFGLTYGGYWTLPOLLETPARFLIFVPDSPARTTF
FMVYLIAFMKPKNALDELALAILVLVRKYLVNAVNMNFVIALVTGDLPEWGMLIASIFF
EAADVGVYSNPFRSLSPFWHALIAAWTLHANDVIDYLFDMMPOYSMLS DYMEIGEYGF
WTSTSIAlAYLVVS KKOTULELM"

gene

CDS

5628..6422
/gene="yjbB"
5628..6422
/gene="yjbB"
/function="hypothetical"
/note="putative"
/citation={2}
/codon_start=1
/transl_table=1
/protein_id="AAB38439.1"
/db_xref="GI:1146231"
/translation="MKRKLTICLLAIIFPNGNAKAARGSLSELINDLSDFVPMTRQQA
ACYEBAQLQLRYEKTIKSAREKKODPMLTGAOIROIILTGNDAWSLKWDPSDIOKR
LRPAAOFRMLMAVNRSDPLNGSLEKYMIAFTELKRDVOYNQSSTSHREKMNFSLIS
YHLIXPSLITDIYESDOLETNVGKHIDIYIQEEEFQQMTSETKLERUSLJQHDIKANFDVY
EBDDADPSLHWIIDTGTSIIITALTYVGRKKYKRNKLKKRDYPK"

terminator

6434..6447
/standard_name="transcription terminator"
/note="putative"
/citation={2}

gene

complement(6464..7111)

Query Match 100.0%; Score 494; DB 1; Length 23775;
Best Local Similarity 100.0%; Pred. No. 3,6e-123;
Matches 494; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

/gene="ypjC"
complement(6464. .7111)
/gene="ypjC"
/function="hypothetical"
/note="putative"

Query 1 GCTAATGTTGGTGTACAAAGCCCGTTGATTTTGGTATACCTTCATTGGGCAATACGCTG 60
Db 13043 GCTAATGTTGGTGTACAAAGCCCGTTGATTTTGGTATACCTTCATTGGGCAATACGCTG 1310
Query 61 CGAATCGACCTATTATTAATAATGATGACATGTGACGACGCTGCGCTGATCCAAAAA 120
Db 13103 CGAATCGACCTATTATTAATAATGATGACATGTGACGACGCTGCGCTGATCCAAAAA 1316
Query 121 GGACTGGGACAGAGGAGTGAATACCTCGCGAATCTTAGAAGTGAAGATCCTTCGTTG 180
Db 13163 GGACTGGGACAGAGGAGTGAATACCTCGCGAATCTTAGAAGTGAAGATCCTTCGTTG 1322
Query 181 TAAAGGAAGTCTTTTGGCTTGAGAAAGAAACGCGAGATCATCTCTTAAACATTAGG 240
Db 13223 TAAAGGAAGTCTTTTGGCTTGAGAAAGAAACGCGAGATCATCTCTTAAACATTAGG 1328
Query 241 AGAGAAAAACATGAAAAACAAAACCTGATTTTCTAAAAATGAAGAGCTGAAGAACGAT 300
Db 13283 AGAGAAAAACATGAAAAACAAAACCTGATTTTCTAAAAATGAAGAGCTGAAGAACGAT 1333
Query 301 TGTACATGCTGACCGCTTATGATTAATCCGCGACGTAACTTGCCTGAACAAGCGGAGTTGA 360
Db 13343 TGTACATGCTGACCGCTTATGATTAATCCGCGACGTAACTTGCCTGAACAAGCGGAGTTGA 1340
Query 361 CATGATTTTAACTCGGTGATTCACCTTGAATGTCCTCGGCTTGAATTCACTGTGG 420
Db 13403 CATGATTTTAACTCGGTGATTCACCTTGAATGTCCTCGGCTTGAATTCACTGTGG 1346
Query 421 TGTGCAAGTTGGCGACATGATTCATCAACAAAAGCCGTTAAAGGGGTGGCGCGAATAC 480
Db 13463 TGTGCAAGTTGGCGACATGATTCATCAACAAAAGCCGTTAAAGGGGTGGCGCGAATAC 1352
Query 481 CTTTATTTGACAG 494
Db 13523 CTTTATTTGACAG 13536

RESULT 3
BSUB0012/c 201375 bp DNA linear BCT 07-JUL-2005
LOCUS BSub0012 subcellis complete genome (section 12 of 21) : from 2207806
to 2409180.
ACCESSION 299115 AL009126
VERSION 299115.2 GI:32468778
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

Bacillus subtilis subsp. subtilis str. 168
Bacillus subtilis subsp. subtilis str. 168
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
1 (bases 1 to 201375)
Kunst, F., Ogasawara, N., Moszer, I., Albertini, A.M., Alloni, G.,
Azevedo, V., Berrtero, M.G., Bessières, P., Bolotin, A., Borchert, S.,
Borriis, R., Boursier, L., Brans, A., Braun, M., Brignell, S.C.,
Bron, S., Brouillet, S., Brusch, C.V., Caldwell, B., Capuano, V.,
Carter, N.M., Choi, S.K., Codani, J.J., Connerston, I.F., Cummings, N.J.,
Daniel, R.A., Denizot, F., Devine, K.M., Dusterhoft, A., Ehrlich, S.D.,
Emmerson, P.T., Entian, K.D., Errington, J., Fabret, C., Fexari, E.,
Fouger, D., Fritz, C., Fujita, K., Fujita, Y., Funai, S., Galizli, A.,
Gallier, D., Ghim, S.Y., Glaser, P., Goffeau, A., Golightly, E.J.,
Graud, G., Guiseppl, G., Guy, B.J., Haga, K., Hales, J., Hazwood, C.R.,
Henaut, A., Hilbert, H., Holstapel, S., Hosono, S., Hullo, M.F.,
Itaya, M., Jones, L., Joris, B., Karamata, D., Kasahara, Y.,
Klaer, Blanchard, M., Klein, C., Kobayashi, Y., Koester, P.,
Kongesteh, G., Krogh, S., Kumano, M., Kurita, K., Lapidus, A.,

Lardinois,S., Lauber,J., Lazarevic,V., Lee,S.M., Levine,A., Liu,H.,
Mairuda,S., Manel,C., Medigue,C., Medina,N., Mellado,R.P.,
Minuro,M., Moestl,D., Nakai,S., Noback,M., Noone,D., O'Reilly,M.,
Ogawa,K., Ogihara,A., Oudega,B., Park,S.H., Patro,V., Pohl,T.M.,
Portetelle,D., Porwollik,S., Prescott,A.M., Prebecan,E., Pulic,P.,
Purnelle,B., Rapoport,G., Rey,M., Reynolds,S., Rieger,M.,
Rivolta,C., Rochae,E., Roche,B., Rose,M., Sadate,Y., Sato,T.,
Scanlan,E., Schleich,S., Schroeter,R., Scoffone,F., Sekiguchi,J.,
Sekowaka,A., Seror,S.J., Serror,P., Shin,B.S., Soldo,B.,
Sotokin,A., Tacconi,E., Takagi,T., Takahashi,H., Takemaru,K.,
Takeuchi,M., Tamakoshi,A., Tanaka,T., Terpetra,P., Tognoni,A.,
Totoato,V., Uchiyama,S., Vandenbol,M., Vannier,F., Vassartotti,A.,
Viari,A., Wambut,R., Wedler,E., Wedler,H., Weltzenegger,T.,
Winters,P., Wipac,A., Yamamoto,H., Yamae,K., Yasumoto,K., Yata,K.,
Yoshida,K., Yoshikawa,H.F., Zumstein,E., Yoshikawa,H. and
Danchin,A.

The complete genome sequence of the gram-positive bacterium
Bacillus subtilis
Nature 390 (6657), 249-256 (1997)

9384377

2 (bases 1 to 201375)
Kunst,F., Ogasawara,N., Yoshikawa,H. and Danchin,A.
Direct Submission
Submitted (27-JUN-2003) I. Moszer, A. Danchin, Institut Pasteur,
Genetique des Genomes Bacteriens, 28 rue du Docteur Roux, 75724
Paris Cedex 15, FRANCE. E-mail: moszer@pasteur.fr
adanchin@pasteur.fr Phone: +33 (0)1 45 68 84 41, Fax: +33 (0)1 45
68 89 48

On Jul 7, 2003 this sequence version replaced gi:2634478.
This entry contains data from release R16.1 of the Subtilist
database. Further data on gene annotation and detailed information
about changes from previous releases can be found at
<http://genolist.pasteur.fr/Subtilist/>.

Location/Qualifiers

1..201375
/organism="Bacillus subtilis subsp. subtilis str. 168"
/mol_type="genomic DNA"
/strain="168"
/db_xref="taxon:224308"
complement(246..371)
/gene="yopL"
/locus_tag="BSU20850"
complement(246..371)
/gene="yopL"
/locus_tag="BSU20850"
/function="unknown"
/codon_start=1
/transl_table=1
/protein_id="CAB14003.1"
/db_xref="GI:2634505"
/db_xref="UniProt/TREMBL:O31926"
/translation="MKKLIMLVILGALGTYSISADSSIQASGDYEVAGMPRA"
complement(385..1545)
/gene="yopK"
/locus_tag="BSU20860"
complement(385..1545)
/gene="yopK"
/locus_tag="BSU20860"
complement(385..1545)
/function="unknown"
/codon_start=1
/transl_table=1
/protein_id="CAB14004.1"
/db_xref="GI:2634506"
/db_xref="UniProt/TREMBL:O31927"
/translation="MELIRIAMKDKLENDNSLMMKMTVAGLKNPNPLYDLNHDGKT
FNEFSIYNIYKSGYQPREYELMDYCLNLTAVTKAASALEYADANFFIEYVLIYD
SMISCSNNKSKYGVYKIHRELSNVITREPAVRGLKINIKPEMNSPRLILLYH
YLTGNSPDMALIKOILDSISENNMYRNTQVRVHYLMSNITKLNISLECEYSK
KALBESTNLRQVRSYITGNSLFSYNYELAQENFLKGLSISVQENYNNMIFQALCE
LNNVRKENKMINFESDISIMDQEAHCFINFNENSKAEVLDKLDLVHNDENLAAH
YLLKRLQNKACFPYSIEYFKSKNDKFLIRLPLLEIQKMGENDKLELLLL"
complement(1722..2138)

gene

/gene="yopJ"
/locus_tag="BSU20870"
complement(1722..2138)
/gene="yopJ"
/locus_tag="BSU20870"
/function="unknown"
/codon_start=1
/transl_table=1
/protein_id="CAB14005.1"
/db_xref="GI:2634507"
/db_xref="UniProt/TREMBL:O31928"
/translation="MKNTFYFRWEDLHGELGVDSFNLRLASYSNLSQQQLVELIKEL
ISIERBILAKFPIHISENAPVDEQHVYKGVAGDMYKMDLLSLVTLALDLTNLLD
HVNIIISLACLSRFPREIPAREPAKDIABERYVSLK"
complement(2140..2673)
/gene="yopI"
/locus_tag="BSU20880"
complement(2140..2673)
/gene="yopI"
/locus_tag="BSU20880"
/function="unknown"
/codon_start=1
/transl_table=1
/protein_id="CAB14006.1"
/db_xref="GI:2634508"
/db_xref="UniProt/TREMBL:O31929"
/translation="MNNIGELISFEGIGALLGVLTLLIHLHLKHFGQIKFYIVDF
EYFKTDNDGNGNMNSKDEAKOIEHSGIEIYNGAEIPKVAIEIKCFYKMTNLIY
SVTPDQATTEPAREGYTRDKLPNNINLPSKOIIAIIITLFLNEKETQVKKCNRYVL
EADHDNKMVKVLEGEF"
complement(2700..3236)
/gene="yopH"
/locus_tag="BSU20890"
complement(2700..3236)
/gene="yopH"
/locus_tag="BSU20890"
/function="unknown"
/codon_start=1
/transl_table=1
/protein_id="CAB14007.1"
/db_xref="GI:2634509"
/db_xref="UniProt/TREMBL:O31930"
/translation="MSDQNKSPSWGDIITKLSPKYLLGLAVSGIGLWGNVGLGKT
LGIKDAINSYKLYGLVFLASTFSILSHFIWMLSLGKKNKIDQKSYKQKERTNLIN
RREKOILSPYIPDVRSEVLSITDQAELEHLKIYRSSNISNRGSYFAYNIQPMAR
GYLTKNKLHBSIDHIL"
complement(3275..3406)
/gene="yopG"
/locus_tag="BSU20900"
complement(3275..3406)
/gene="yopG"
/locus_tag="BSU20900"
/function="unknown"
/codon_start=1
/transl_table=1
/protein_id="CAB14008.1"
/db_xref="GI:2634510"
/db_xref="UniProt/TREMBL:O31931"
/translation="MYTIEWENGKKNIYABGMIEMAILLEDLYQKRFYEWKKL"
complement(3417..3632)
/gene="yopF"
/locus_tag="BSU20910"
complement(3417..3632)
/gene="yopF"
/locus_tag="BSU20910"
/function="unknown"
/codon_start=1
/transl_table=1
/protein_id="CAB14009.1"
/db_xref="GI:2634511"
/db_xref="UniProt/TREMBL:O31932"
/translation="MPLIDYFYVLQENKKEYFAAFLKLDSEGYLTSSDLHBAASKIHNNL
EVIYVASELTKKNVQCEVREIQVYK"
complement(1722..2138)

gene

```

gene      complement (3636. .3887)
          /gene="yopB"
          /locus_tag="BSU20920"
          complement (3636. .3887)
          /gene="yopB"
          /locus_tag="BSU20920"
          /function="unknown"
          /codon_start=1
          /transl_table=11
          /protein_id="CAB14010.1"
          /db_xref="GI:2634512"
          /db_xref="UniProt/TREMBL:O31933"
          /translation="MIGLAFLLIWLGVGLTGKIFVVDQVYDDEEKLMDKETAAG
          MERTLASLFFKNKLNVIAFEMLIGLPLAMKITLPRG"
          complement (4474. .4869)
          /gene="yopD"
          /locus_tag="BSU20930"
          complement (4474. .4869)
          /gene="yopD"
          /locus_tag="BSU20930"
          /function="unknown"
          /locus_tag="BSU20930"
          /note="similar to unknown proteins from B. subtilis"
          /codon_start=1
          /transl_table=11
          /protein_id="CAB14011.1"
          /db_xref="GI:2634513"
          /db_xref="UniProt/TREMBL:O31934"
          /translation="MSLNOKALNSVYVTLAFSLFGLIFGLYLVYSGFMAIALVTI
          AIIAPVLLITVLVPAAPLOVWLRRRRRFSFLINFLIYIAVAFSAVLFVFWFDYDPNML
          TMRSPFYINISIVAAPLYFWDSIFLRN"

Query Match      100.0%; Score 494; DB 1; Length 201375;
Best Local Similarity 100.0%; Pred. No. 3.3e-123;
Matches 494; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CDS
1      GCTAATGTTGGTGCACAGCCCGTGGATTTGGTACTCTTCATTCGCGCAGATCCGCTG 60
Db      146314 GCTAATGTTGGTGCACAGCCCGTGGATTTGGTACTCTTCATTCGCGCAGATCCGCTG 146255
Qy      61      GGAAGTGCACCTATATTAATAATAGATAGACATTCGACAGCTGCGCTTGATCCAAAAA 120
Db      146254 GGAAGTGCACCTATATTAATAATAGATAGACATTCGACAGCTGCGCTTGATCCAAAAA 146195
Qy      121     GGAAGTGCACAGAGGATGAAATCGCCGAACTTTAGAAAGTGAAGAAATCCTTCGTTG 180
Db      146194 GGAAGTGCACAGAGGATGAAATCGCCGAACTTTAGAAAGTGAAGAAATCCTTCGTTG 146135
Qy      181     TAACGGAAGCTTTTGGCTTCGACAGAAAAAGCGCAATCATCTCCCTTAAACATGAGG 240
Db      146134 TAACGGAAGCTTTTGGCTTCGACAGAAAAAGCGCAATCATCTCCCTTAAACATGAGG 146075
Qy      241     AGGAGAAAACATGAAAACAAACCTGATTTCTAAAAATGAAGAGCTGAAAGAACCAT 300
Db      146074 AGGAGAAAACATGAAAACAAACCTGATTTCTAAAAATGAAGAGCTGAAAGAACCAT 146015
Qy      301     TGTCAATGCTGACCGCTTATGATTATCCGCAAGCTAACTTGTCTGAACAAGCGGAGTTGA 360
Db      146014 TGTCAATGCTGACCGCTTATGATTATCCGCAAGCTAACTTGTCTGAACAAGCGGAGTTGA 145955
Qy      361     CATGATTTTATGTCGGTATTCCTTGAATGCTGCTCGCGCTTGATTAACCTGTGG 420
Db      145954 CATGATTTTATGTCGGTATTCCTTGAATGCTGCTCGCGCTTGATTAACCTGTGG 145895
Qy      421     TGTGACAGTTGGCGGACATGATCATATCAAAAAGCGTTAAAGGGGTGCGCGAATAC 480
Db      145894 TGTGACAGTTGGCGGACATGATCATATCAAAAAGCGTTAAAGGGGTGCGCGAATAC 145835
Qy      481     CTTTATTTGTGACAG 494
Db      145834 CTTTATTTGTGACAG 145821

```

```

AX100815
LOCUS      AX100815      2363 bp      DNA      linear      PAT 10-APR-2001
DEFINITION Sequence 59 from Patent WO0121772.
ACCESSION AX100815
VERSION    AX100815.1 GI:13619747
KEYWORDS
SOURCE
ORGANISM   Bacillus subtilis
            Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE
  1. Yocum,R.R., Patterson,T.A., Hermann,T. and Petro,J.G.
  Methods and microorganisms for production of panto-compounds
  Patent: WO 0121772-A 59 29-MAR-2001;
  OMNIGENE BIOPRODUCTS, INC. (US)
  Location/Qualifiers
    1..2363
    /organism="Bacillus subtilis"
    /mol_type="unassigned DNA"
    /db_xref="taxon:1423"
    242..1075
    /note="unnamed protein product"
    /codon_start=1
    /transl_table=11
    /protein_id="CAC36485.1"
    /db_xref="GI:13619748"
    /translation="MKTKDLPLMKSESEPIVWLTVADYPAKLAQAGVDMILVDS
    LGWVGLDSTVGYTADMTHTKAYKRGAPNPIYDMPMSHLSKETTLKRAAI
    VQSGDALKLEGEVSESIRLVLTGSLIVSHLSLTPQSVGLGTYKQGDQQA
    KKLIESIKCEBGAAMLVLECVAPLTAIAETLSIPVIGIAGVADQQLVYTHI
    IGHGERTPKFVKQYTRIDETIETALSISGYQDVNRHAFPRQKHSFQNRQTVLDL
    YGK"
    1077..1937
    /note="unnamed protein product"
    /codon_start=1
    /transl_table=11
    /protein_id="CAC36486.1"
    /db_xref="GI:13619749"
    /translation="MRQITDISQKEAIIKQHSRKSIGFVPTMGFLHGLTLADKA
    RQENDAVIMSIPVNPAPQFQNEDEAPRIERDALALBNAGVDILTPPAHMPRE
    KNVNIHERETDVLGSRSGHEDGVAIVLTKLNLVKTFRATFGLDQQAIVVDGL
    ISDFMDIELVADTVREBDGLAKSSNNVLTREKRAPRLYRALQTSLELVQGER
    DPEAVIKAKDIIETTSGLIDVELYSYPELEPVNEIAGKMLLAVAFAFSKRLDNI
    IIDIREMER"
    1939..2322
    /note="unnamed protein product"
    /codon_start=1
    /transl_table=11
    /protein_id="CAC36487.1"
    /db_xref="GI:13619750"
    /translation="MYRTMSGKLHRAVTVEANLVYGSITIDEDLIDAVGMLPNRY
    QIVNNNGARLEIYIIPGKSGSVTCANGAARLVQSGDKVIITISYKMSDQEAASHE
    PKVAVLNDQNKIEQMLGNEPARTL"

ORIGIN
Query Match      98.2%; Score 485; DB 6; Length 2363;
Best Local Similarity 100.0%; Pred. No. 1.1e-120; Indels 0; Gaps 0;
Matches 485; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CDS
10      TTGGTACAAGCCCGTGGATTTTGGTATCTTCATTTGGGACAGTATCGCTCGAATGCA 69
Db      1      TTGGTACAAGCCCGTGGATTTTGGTATCTTCATTTGGGACAGTATCGCTCGAATGCA 60
Qy      70      CCTATTATTAATAGATGACATTCGACAGCTGCTGCTTGATCCAAAAAGAGCTGGGA 129
Db      61      CCTATTATTAATAGATGACATTCGACAGCTGCTGCTTGATCCAAAAAGAGCTGGGA 120
Qy      130     CAGAGGATGAAATCCGCCGAACTTTGAAGTGAAGAAATCCTCTGTTGAACGGAAG 189
Db      121     CAGAGGATGAAATCCGCCGAACTTTGAAGTGAAGAAATCCTCTGTTGAACGGAAG 180
Qy      190     GTTTTGGCTTGCAAGAAAACGCGAGATCATCTCTCTAAACATGAGGAGAGAAA 249
Db      181     GTTTTGGCTTGCAAGAAAACGCGAGATCATCTCTCTAAACATGAGGAGAGAAA 240

```

RESULT 4

LOCUS AX572874 6725 bp DNA linear PAT 29-NOV-2002
DEFINITION Sequence 22 from Patent WO02057476.
ACCESSION AX572874
VERSION AX572874.1 GI:26004960
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 Hermann, T., Patterson, T.A., Pero, J.G., Yocum, R.R., Baldeusius, K.U. and Beck, C. Methods and microorganisms for the production of 3-(2-hydroxy-3-methyl-butylamino)-propionic acid (hmbpa) Patent: WO 02057476-A 22 25 -JUL-2002;
JOURNAL Omnigene Bioproducts (US)
FEATURES
LOCATION/Qualifiers
SOURCE
1. .6725
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="vector-pam636"

Db 250 CATGAAAACAAACTGATTTTCTTAAATGAAAGAGCTGAAAGAACGATTGTCATGCT
241 CATGAAAACAAACTGATTTTCTTAAATGAAAGAGCTGAAAGAACGATTGTCATGCT
Qy 310 GACCGCTTATGATTAATCCGAGAGCTAAACCTTGCTGAACAGCGGAGTTGACATGATTTT
Db 301 GACCGCTTATGATTAATCCGAGAGCTAAACCTTGCTGAACAGCGGAGTTGACATGATTTT
Qy 370 AGTCGTATTCATCTTGAAATGCTGTCTCGGCTTATTCACCTGTCGCTGACAGT 429
Db 361 AGTCGTATTCATCTTGAAATGCTGTCTCGGCTTATTCACCTGTCGCTGACAGT 420
Qy 430 TGGCGACATGATTCATCAACAAAGCCGTTAAAGGGGTCGCGCAATACCTTATTTGT 489
Db 421 TGGCGACATGATTCATCAACAAAGCCGTTAAAGGGGTCGCGCAATACCTTATTTGT 480
Qy 490 GACAG 494
Db 481 GACAG 485

RESULT 5
AX100850/c 7381 bp DNA linear PAT 10-APR-2001
LOCUS AX100850
DEFINITION Sequence 94 from Patent WO0121772.
ACCESSION AX100850
VERSION AX100850.1 GI:13619778
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 Yocum, R.R., Patterson, T.A., Hermann, T. and Pero, J.G. Methods and microorganisms for production of panto-compounds Patent: WO 0121772-A 94 29-MAR-2001;
JOURNAL OMNIGENE BIOPRODUCTS, INC. (US)
FEATURES
LOCATION/Qualifiers
SOURCE
1. .7381
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Description of Artificial Sequence: recombinant PAN06 plasmid"

ORIGIN
Query Match 52.2%; Score 258; DB 6; Length 7381;
Best Local Similarity 100.0%; Pred. No. 4.5e-59;
Matches 258; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 237 GAGGAGGAGAAACATGAAACAAACATGATTTTCTTAAATGAAAGAGCTGAAAGAAC 296
Db 3652 GAGGAGGAGAAACATGAAACAAACATGATTTTCTTAAATGAAAGAGCTGAAAGAAC 3593
Qy 297 CGATTGTCATGCTGACCGCTTATGATTAATCCGAGAGCTAAACCTTGCTGAACAGCGGAG 356
Db 3592 CGATTGTCATGCTGACCGCTTATGATTAATCCGAGAGCTAAACCTTGCTGAACAGCGGAG 3533
Qy 357 TTGACATGATTTTATGCTGATTTGATCTTGGAATGCTGCTCGGCTTATTCATCTG 416
Db 3532 TTGACATGATTTTATGCTGATTTGATCTTGGAATGCTGCTCGGCTTATTCATCTG 3473
Qy 417 TCGGTGACAGTTGTCGAGCATGATCCATCAACAAAGCCGTTAAAGGGGTCGCGCA 476
Db 3472 TCGGTGACAGTTGTCGAGCATGATCCATCAACAAAGCCGTTAAAGGGGTCGCGCA 3413
Qy 477 ATACCTTATTTGTGACAG 494
Db 3412 ATACCTTATTTGTGACAG 3395

RESULT 6
AX572874

LOCUS AX572874 6725 bp DNA linear PAT 29-NOV-2002
DEFINITION Sequence 22 from Patent WO02057476.
ACCESSION AX572874
VERSION AX572874.1 GI:26004960
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 Hermann, T., Patterson, T.A., Pero, J.G., Yocum, R.R., Baldeusius, K.U. and Beck, C. Methods and microorganisms for the production of 3-(2-hydroxy-3-methyl-butylamino)-propionic acid (hmbpa) Patent: WO 02057476-A 22 25 -JUL-2002;
JOURNAL Omnigene Bioproducts (US)
FEATURES
LOCATION/Qualifiers
SOURCE
1. .6725
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="vector-pam636"

ORIGIN
Query Match 52.0%; Score 257; DB 6; Length 6725;
Best Local Similarity 100.0%; Pred. No. 8.5e-59;
Matches 257; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 238 AGGAGGAGAAACATGAAACAAACATGATTTTCTTAAATGAAAGAGCTGAAAGAAC 297
Db 325 AGGAGGAGAAACATGAAACAAACATGATTTTCTTAAATGAAAGAGCTGAAAGAAC 384
Qy 298 GATTGTCATGCTGACCGCTTATGATTAATCCGAGAGCTAAACCTTGCTGAACAGCGGAGT 357
Db 385 GATTGTCATGCTGACCGCTTATGATTAATCCGAGAGCTAAACCTTGCTGAACAGCGGAGT 444
Qy 358 TGACATGATTTTATGCTGATTTGATCTTGGAATGCTGCTCGGCTTATTCATCTG 417
Db 445 TGACATGATTTTATGCTGATTTGATCTTGGAATGCTGCTCGGCTTATTCATCTG 504
Qy 418 CGGTGACAGTTGTCGAGCATGATCCATCAACAAAGCCGTTAAAGGGGTCGCGCA 477
Db 505 CGGTGACAGTTGTCGAGCATGATCCATCAACAAAGCCGTTAAAGGGGTCGCGCA 564
Qy 478 TACCTTATTTGTGACAG 494
Db 565 TACCTTATTTGTGACAG 581

RESULT 7
AX573226 6725 bp DNA linear PAT 29-NOV-2002
LOCUS AX573226
DEFINITION Sequence 22 from Patent WO02057474.
ACCESSION AX573226
VERSION AX573226.1 GI:26005139
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 Hermann, T., Patterson, T.A., Pero, J.G., Yocum, R.R., Baldeusius, K.U. and Beck, C. Processes for enhanced production of pantothenate Patent: WO 02057474-A 22 25 -JUL-2002;
JOURNAL OMNIGENE BIOPRODUCTS, INC. (US)
FEATURES
LOCATION/Qualifiers
SOURCE
1. .6725
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="vector-pam636"

ORIGIN
Query Match 52.0%; Score 257; DB 6; Length 6725;

RESULT 9

/db_xref="GI:13619712"

/translation="MKTKLDFLKKKSEEPVLMTAYDYPAPAKLAEQGVMLTVGS
LGMVVLGDSVTGTVADMHTKAVKRGAPNTHLYMDMPMSVHLSEKDTLKNAAAI
VQESGADALKEEGSGVPESEIRALTLGGIPIVSHLGLTPQSVGLGKYGKQBSA
KKLIEDSTKESAGMMVLHCVPALPTAKIAETLSIPVIGAGVAKADGVLYHDI
IGHVERTPKFKQYTRIDETIETASGVODVHRAPPEQKHSFQMNQTVLDGLYGG
K"

ORIGIN

Query Match 49.4%; Score 244; DB 6; Length 811;

Best Local Similarity 100.0%; Pred. No. 3.2e-55;

Matches 244; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 251 ATGAAACAAAACGATGTTTCTAAATGAAGAGCTGAAGAACCGATTGTCATGCTG 310
Db 1 ATGAAACAAAACGATGTTTCTAAATGAAGAGCTGAAGAACCGATTGTCATGCTG 60
QY 311 ACCGCTTATGATTATCCGCGACGCTAACTTGCTGAACAACGGGAGTTGACATGATTTTA 370
Db 61 ACCGCTTATGATTATCCGCGACGCTAACTTGCTGAACAACGGGAGTTGACATGATTTTA 120
QY 371 GTCGATGTTCACTTGGAAATGCTGCTCGGCTTGAATTCATCTGTGCTGTGACAGTT 430
Db 121 GTCGATGTTCACTTGGAAATGCTGCTCGGCTTGAATTCATCTGTGCTGTGACAGTT 180
QY 431 GCGGACATGATCCATCATACAAAGCCGTTAAAGGGGTCGCGGATACCTTTATTTG 490
Db 181 GCGGACATGATCCATCATACAAAGCCGTTAAAGGGGTCGCGGATACCTTTATTTG 240
QY 491 ACAG 494
Db 241 ACAG 244

RESULT 11
AE017333_23/c

WPCOMMENT

Sequence split into 43 fragments LOCUS AE017333 Accession AE017333

Fragment Name	Begin	End
AE017333_00	1	110000
AE017333_01	100001	210000
AE017333_02	200001	310000
AE017333_03	300001	410000
AE017333_04	400001	510000
AE017333_05	500001	610000
AE017333_06	600001	710000
AE017333_07	700001	810000
AE017333_08	800001	910000
AE017333_09	900001	1010000
AE017333_10	1000001	1110000
AE017333_11	1100001	1210000
AE017333_12	1200001	1310000
AE017333_13	1300001	1410000
AE017333_14	1400001	1510000
AE017333_15	1500001	1610000
AE017333_16	1600001	1710000
AE017333_17	1700001	1810000
AE017333_18	1800001	1910000
AE017333_19	1900001	2010000
AE017333_20	2000001	2110000
AE017333_21	2100001	2210000
AE017333_22	2200001	2310000
AE017333_23	2300001	2410000
AE017333_24	2400001	2510000
AE017333_25	2500001	2610000
AE017333_26	2600001	2710000
AE017333_27	2700001	2810000
AE017333_28	2800001	2910000
AE017333_29	2900001	3010000
AE017333_30	3000001	3110000
AE017333_31	3100001	3210000
AE017333_32	3200001	3310000
AE017333_33	3300001	3410000
AE017333_34	3400001	3510000

AE017333_35 3500001 3610000
AE017333_36 3600001 3710000
AE017333_37 3700001 3810000
AE017333_38 3800001 3910000
AE017333_39 3900001 4010000
AE017333_40 4000001 4110000
AE017333_41 4100001 4210000
AE017333_42 4200001 4222645

Continuation (24 of 43) of AE017333 from base 2300001 (AE017333 Bacillus licheniformis DK

Query Match 46.4%; Score 229; DB 1; Length 110000;

Best Local Similarity 73.3%; Pred. No. 3e-51;

Matches 366; Conservative 0; Mismatches 115; Indels 18; Gaps 5;

QY 9 GTTGATCAAGCCCGTGAATTTTGATTAATTCATTCGTCGACATGCTGCGAATGCG 68
Db 25528 GTTGATCAAGATCCATGCGGTTTGGTATTAATTCGATGCGGACATGCTTCGGAATGCG 25469
QY 69 ACTTATTAATTAATATAGATGACATTCGACAGTCTGCTTGAATCCAAAAAGACTGCG 128
Db 25468 ACTTATTAATTAATAGATGACATTCGACAGTCTGCTTGAATCC--AAAAAGACTGCG 25411
QY 129 ACAGAGGATGAACCTGCCGAACCTTTAGAAATGAAGATCCTTCTGTTAAGGAA 188
Db 25410 ACAGAGGATGAATCTTCCGGAAC-ACGAGATAGGAATCCTTCTCTTG--CCGGA 25354
QY 189 GGTGTTTGGCTTGCAGAAAGAA-----ACGCGATGATCCTCCTTAACA 235
Db 25353 GGTGTTTGGCTTGCAGAAAGATGCGATGCAAGCAGGACCTCATCTCTGAAAA 25294
QY 236 TGAGAGAGAGAAAAATGAAGAAAACTGATTTTCTAAAAATGAAGAGTGTGAAGAA 295
Db 25293 AGAGAGAGAGAAAAATGAAGAAAACTGATTTTCTAAAAATGAAGAGAGAG 25234
QY 296 CCGATTGTCACTGTCAGCCTTATGATTAATTCGCGACGCTAACTTGTGAACAAGCGGGA 355
Db 25233 CCGATTGTCACTGTCAGCCTTATGATTAATTCGCGACGCTAACTTGTGTGAACAAGCGGCG 25174
QY 356 GTTGACATGAATTTTATGTCGCTGATTCATCTTGAATGCTGCTCGGCTTGAATTA 415
Db 25173 GTTGACATGAATTTTATGTCGCTGATTCATCTTGAATGCTGCTCGGCTTGAATTA 25114
QY 416 GTCGATGACAGTTCGCGACATGATTCATACAAAGCCGTTAAAGGGTGCAGCG 475
Db 25113 GTCGATGACAGTTCGCGACATGATTCATACAAAGCCGTTAAAGGGTGCAGCG 25054
QY 476 AATACCTTAATGTCAGAG 494
Db 25053 AATACCTTAATGTCAGAG 25035

RESULT 12

CP000002_23/c

WPCOMMENT

Sequence split into 43 fragments LOCUS CP000002 Accession CP000002

Fragment Name	Begin	End
CP000002_00	1	110000
CP000002_01	100001	210000
CP000002_02	200001	310000
CP000002_03	300001	410000
CP000002_04	400001	510000
CP000002_05	500001	610000
CP000002_06	600001	710000
CP000002_07	700001	810000
CP000002_08	800001	910000
CP000002_09	900001	1010000
CP000002_10	1000001	1110000
CP000002_11	1100001	1210000
CP000002_12	1200001	1310000
CP000002_13	1300001	1410000
CP000002_14	1400001	1510000
CP000002_15	1500001	1610000
CP000002_16	1600001	1710000

QY	61	CGAACTGCACCTATTATTAAATATGATATGACATTTGCAGCACTGTCCTTATCCAAAAA	120
Db	2849	CGAACTGCACCTATTATTAAATATGATATGACATTTGCAGCACTGTCCTTATCCAAAAA	2790
QY	121	GGACTGGGACAGAGGATGAATACTGCCGCACTTAGAAAGTGAAGATCCTTCCTGTG	180
Db	2789	GGACTGGGACAGAGGATGAATACTGCCGCACTTAGAAAGTGAAGATCCTTCCTGTG	2730
QY	181	TAAACGAAGGTTTTTTGGCTTGCAGAGAAAAACGGCAGATC	221
Db	2729	TAAACGAAGGTTTTTTGGCTTGCAGAGAAAAACGGCAGATC	2689
RESULT 15			
BACJ0JC			
LOCUS			
DEFINITION	BACJ0JC	6540 bp	DNA linear BCT 26-MAR-1996
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
COMMENT			
FEATURES			
source			
terminator			
gene			
CDS			
gene			
CDS			

gene
CDS
/codon_start=1
/transl_table=1
/product="unknown"
/protein_id="AA92873.1"
/db_xref="GI:755602"
/translation="MSDKTMKDIOAEVDRIYIGPKESGFSPPLAMMARLTSEGLAE
VNHVGEKPKATDDKSMEEIGDVLFLVCLANSIDISEAHDRVMHKFTLRDD
RWTRKEBK"
1358..2101
/gene="dapB"
/feature="lysine biosynthesis"
/note="28.2% of identical aminoacids with the Escherichia
coli dihydrodipicolinate reductase; putative"
/codon_start=1
/transl_table=1
/product="dihydrodipicolinate reductase"
/protein_id="AA92874.1"
/db_xref="GI:755603"
/translation="MSNETIKLVLAGRSGMGEAVKLAERTPHFDLVGAIIDHTYDQO
KLSDMPVPSADFAITYDTHACFTETQPDVAILDTPPIGKVTAKIALSEHVRPVGTT
GSEFDELKELTEEGKIGALIANPNEGLIMMKRSKMAAYFEVEVIELHDDK
LDAEGSTLKTAEMLSEVRKEKQGGHDEKILPGANGAONGIRLRLRIDRDSYNR
ASPSMGVLSYEQVMKIDQLVYGLNIIID"
2117..2530
/gene="jojF"
2117..2530
/gene="jojF"
/note="putative"
/codon_start=1
/transl_table=1
/product="unknown"
/protein_id="AA92875.1"
/db_xref="GI:755604"
/translation="MKIALIADKKKODMVOPTTAVRDIILKNHDLVATGTGKIHRA
TGLQIERQSGPFLGGDOQIGALIANALDVLFIPLRDLTPQHPHPPVSAIRLCDVYS
IPLANNMGTAEILVRTLDEGVFEFRDLLRGEPPNV"
2523..3197
/gene="jojG"
2523..3197
/gene="jojG"
/note="putative"
/codon_start=1
/transl_table=1
/product="unknown"
/protein_id="AA92876.1"
/db_xref="GI:755605"
/translation="MKNADVIAFGASHDDVEIGMGITAKFVKOEKVMICDLTEAE
LSNNGVSRKPEBAEARIILGADRIGQLTPEKGLINSIDAIRISIVVILCRKAFV
MPKYSRDKRKEAALAVEEALIFAGILHDEKESLPHNHSKVVYIMINGFQPDV
IDISITIAKKQSLNAYKSGFIPSKDSVSTPLTNGYIEVAREKIJLKGAGVEYAEV
SPNNGC"
3229..4302
/gene="jojH"
3229..4302
/gene="jojH"
/note="21.6% of identity to trans-acting transcription
factor of Saccharomyces cerevisiae; 25% identity to sucrose
synthase of Zea mays; putative"
/codon_start=1
/transl_table=1
/product="unknown"
/protein_id="AA92877.1"
/db_xref="GI:755606"
/translation="WRKLGKQLAEKGEIHFTSSIPRLNTYHPNTHHEVEVNOYA
VFCPPYDULATASKIAIEVAERENDIITHAHVAPHAVCAVLAQMLKRNIGVITTLG
TDITVLGDPDLKDLIPALFESDRTVAVSALAAETVILKEPKETIETYNFIDEEV
YLKNTAIKIEKHGILPDEKVVIVHSNRPKRVQVYIRVBNINAGTKAKLLVGGQV
PEKRNAGCLIRKGVLEDOVYMLGNQDVEDLYSIDLKLSEKSGVLLLEMACG
VPCIGTNGIGIPETVIAKNVSGFLVVDGVDVTAATRAAMSILDEQLSNRFTKAIEMLE
NEFSSKIVSVEQDIYADLAEPE"
4307..5440
gene

CDS

```

/gene="paps"
4307..5440
/gene="paps"
/note="38.9% of identical aminoacids with the Escherichia
coli pcmb gene product, encoding poly(A) polymerase;
putative"
/codon_start=1
/transl_table=11
/product="polyA polymerase"
/protein_id="AA92878.1"
/db_xref="gi:755607"
/translation="MEKVFIALPVLNLIIEAGHOAVFGAVRDSYMKRTIGDVIA
TDAPDOVERLFOFTVYGDSESDYDPRRSEYQFISLEBDRDLITINAMATA
DGKLDYFGGKDIIDKQVIRTVGKEDRPODALRMLRAYPMGOLGPTLSPETEBAI
AKESKLSHVSVERKTIIEFKLQGRASROALQTLIOTRLYEELPGYHRENLSTS
EPPFSLTSREELMAALLINGLVLDKAPLFLKMKLPKVKIEAHIADTFQSLDA
MTWYRAGKKALISAQISQLRNENKDEKDKDIQYAYQNLPIKSLKDLDTGKDLA
LNRPRAGWVSEELQWIEQAVVTGKLSNKKHIEEWLTCGQH"
5425..6342
/gene="bira"
5425..6342
/gene="bira"
/function="bifunctional biotin operon repressor"
/note="26.7% identity to the Escherichia coli bifunctional
biotin operon repressor"
/codon_start=1
/evidence=experimental
/transl_table=11
/product="biotin [acetyl-CoA carboxylase] ligase"
/protein_id="AA92879.1"
/db_xref="gi:755608"
/translation="MRSTLRKDLIELFSQAGNERISGKISDALGCSRVAWKIEEL
RKEGYEAVRKRGLIKKPKLSESIIRFGLTQVWGQHLIYHVLSTQKTAHEL
ANNNAPBETLENGWMSLILRPDIPLQKTPQLTLAAVAVQGIIEAAGIQTDIKP
NDILNGKTVGILTEMOABEDRVSVIIGINWQOPNDPDELKDIAATLSQAAQ
EKIDRAGVIOHILCFEKKRYDMVTHGFTPIKLWESYALGIGTMMRARTLNGTFYQK
ALGIDDEGVLLTNEGIKIKYSADIEIG"
6501..6533
/note="putative"

```

terminator

```

/note="putative"

```

ORIGIN

```

Query Match      41.1%; Score 203; DB 1; Length 6540;
Best Local Similarity 100.0%; Pred. No. 3.9e-44;
Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  GCTAATGTTGGTACAAGCCGTTGATTGGTATACCTCCATTGGGCAGTATGCCTG 60
      |||
DB      6338 GCTAATGTTGGTACAAGCCGTTGATTGGTATACCTCCATTGGGCAGTATGCCTG 6397

QY      61  CGAAGTGCACCTATTATTAAATAGATAGACATTGCAGCAGTCTGCTTGATCCAAAAA 120
      |||
DB      6398 CGAAGTGCACCTATTATTAAATAGATAGACATTGCAGCAGTCTGCTTGATCCAAAAA 6457

QY      121 GGAATGGGACAGAGGATGAATCTGCCGAACCTTAGAAAGTGAATCCCTTCGTTG 180
      |||
DB      6458 GGAATGGGACAGAGGATGAATCTGCCGAACCTTAGAAAGTGAATCCCTTCGTTG 6517

QY      181 TAAAGGAAGTTTTTTGGCTTGC 203
      |||
DB      6518 TAAAGGAAGTTTTTTGGCTTGC 6540

```

Search completed: September 8, 2005, 21:13:00
 Job time : 2370 secs

This Page Blank (uspic)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 8, 2005, 16:46:49 ; Search time 7320 Seconds
(without alignments)
399.502 Million cell updates/sec

Title: US-09-925-824A-1
Perfect score: 494
Sequence: 1 gctaatgtgtgtgacacg.....gaatacctatgtgacag 494

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N_Geneseq_16Dec04:
1: geneeqn1980s:
2: geneeqn190s:
3: geneeqn200s:
4: geneeqn2001s:
5: geneeqn2001bs:
6: geneeqn2002s:
7: geneeqn2002bs:
8: geneeqn2003s:
9: geneeqn2003bs:
10: geneeqn2003cs:
11: geneeqn2003ds:
12: geneeqn2004s:
13: geneeqn2004bs:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	494	100.0	494	6	ABA95293 B. subtil
2	485	98.2	2363	4	AA502314 B. subtil
3	258	52.2	7381	4	AA502335 Plasmid P
4	257	52.0	6725	6	AA148885 B. subtil
5	257	52.0	6725	6	AA141229 Pantoch
6	257	52.0	8503	4	AA502334 Plasmid P
7	257	52.0	10801	4	AA502333 Plasmid P
8	244	49.4	831	4	AA500988 B. subtil
9	224.6	45.5	10181	12	ADN36367 Plasmid P
10	216.2	43.8	6886	6	AA148883 B. subtil
11	216.2	43.8	6886	6	AA141227 Pantoch
12	151.2	30.6	834	6	ABK73715 Bacillus
13	139	28.1	837	6	ACA21619 Prokaryot
14	137	27.7	1513	6	ABQ70448 Listeria
15	137	27.7	110000	6	ABQ69245-20 Continuation (21 o
16	137	27.7	319630	6	ABQ67194 Listeria
17	133.8	27.1	110000	6	ABA03041_19 Continuation (20 o
18	128.8	26.1	110000	8	ACA36655 Prokaryot
19	117.2	23.7	1880	4	AAH55062 S. epider
20	111.6	22.6	816	8	ACA46771 Prokaryot

21	111.6	22.6	825	6	ABN92106	ABN92106 Staphyloc
22	111.6	22.6	825	13	AD501749	Ad501749 Staphyloc
23	105.8	21.4	865	8	ACA22421	ACA22421 Prokaryot
24	103.2	20.9	792	4	AAH53012	AAH53012 S. epider
25	99.6	20.2	4549	2	AAV74485	AAV74485 Staphyloc
26	94	19.0	787	4	AA552191	AA552191 Staphyloc
27	94	19.0	816	8	ACF74215	ACF74215 Staphyloc
28	92.4	18.7	825	8	ACA20255	ACA20255 Prokaryot
29	87.4	17.7	825	8	ACA32887	ACA32887 Prokaryot
30	87.4	17.7	825	8	ACA33554	ACA33554 Prokaryot
31	87.4	17.7	825	8	ACA33554	ACA33554 Prokaryot
32	87.4	17.7	825	10	ADCC90732	ADCC90732 E. faeciu
33	87.4	17.7	828	10	ADH85238	ADH85238 Enterococ
34	85.6	17.3	792	4	AA554927	AA554927 Staphyloc
35	85.4	17.3	491	4	AA550306	AA550306 Staphyloc
36	85.4	17.3	491	8	ACA17539	ACA17539 Prokaryot
37	85.4	17.3	816	8	ACA47659	ACA47659 Prokaryot
38	78	15.8	3301	4	AAH54845	AAH54845 S. epider
39	74.4	15.1	831	8	ACA27876	ACA27876 Prokaryot
40	74.2	15.0	855	9	ADA32533	ADA32533 DNA encod
41	72.8	14.7	807	8	ACA21162	ACA21162 Prokaryot
42	72	14.6	349980	5	AAH41224	AAH41224 Pyrococcu
43	71.8	14.5	825	8	ACA28486	ACA28486 Prokaryot
44	71.6	14.5	825	8	ACA28877	ACA28877 Prokaryot
45	70	14.2	789	8	ACA41207	ACA41207 Prokaryot

ALIGNMENTS

RESULT 1
ABA95293 standard; DNA, 494 BP.

ABA95293; 10-UN-2002 (first entry)
B. subtilis panB promoter sequence.
Bacterium; tetrahydrofolate; THF; panB; promoter; para-aminobenzoic acid;
PABA; antibacterial; ds.
Bacillus subtilis.
WO200214559-A2.
21-FEB-2002.
10-AUG-2001; 2001WO-US041665.
11-AUG-2000; 2000US-0224925P.
(MILL-) MILLENNIUM PHARM INC.
Murphy C;
WPI; 2002-269209/31.
Identifying inhibitor of bacterial tetrahydrofolate biosynthesis for
treating bacterial infection, by contacting a cell having ketopantoate
PT hydroxymethyltransferase promoter with an agent and measuring promoter
activity.
Claim 3; Page 14-15; 31p; English.
The invention relates to a method of determining whether a test compound
is an inhibitor of bacterial tetrahydrofolate (THF) biosynthesis. The
method involves contacting a bacterial cell with the test compound, where
the cell contains a promoter, the activity of which is increased in the
presence of a compound that inhibits THF biosynthesis, and measuring
activity of the promoter. The method is useful for determining whether a
test compound is an inhibitor of bacterial THF biosynthesis, and also for

determining whether a test compound is an antibacterial agent. The compounds that inhibit THF biosynthesis can be identified by their ability to increase the activity of the B. subtilis panB promoter. The inhibition of THF biosynthesis is detected as inhibition of para-aminobenzoic acid (PABA) uptake into cells and the inhibition is measured in a biochemical assay with a cell extract for an enzyme activity which is required for THF biosynthesis. The inhibitor identified by the method is useful as a lead compound for preparing antibacterial agents. Compositions comprising the antibacterial agent is useful for treating a bacterial infection caused by S. pneumoniae, S. pyogenes, S. agalactiae, S. endocarditis, S. faecium, S. sanguis, S. viridans, S. haemolyticus in a mammal, in particular a human. The compounds can also be used to treat infection of gram negative bacteria e.g., Shigella, E. coli, Klebsiella and Yersinia. The present sequence represents the B. subtilis panB promoter sequence

Sequence 494 BP; 147 A; 95 C; 122 G; 130 T; 0 U; 0 Other;

Query Match 100.0%; Score 494; DB 6; Length 494;
Best Local Similarity 100.0%; Pred. No. 3.6e-143;
Matches 494; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCTATGTTGGTACAAAGCCGTTGATTTGGTATCTTCATGGGCGATGCTCG 60
DB 1 GCTATGTTGGTACAAAGCCGTTGATTTGGTATCTTCATGGGCGATGCTCG 60
OY 61 CGAATCGACCTATTATTAATAGATGACATTCGACGAGTCTGCTGATCCAAAAA 120
DB 61 CGAATCGACCTATTATTAATAGATGACATTCGACGAGTCTGCTGATCCAAAAA 120
OY 121 GGAATGGGACAGAGGATGAAAACGCGAATTTAGAAAGTGAAGAAATCTTCGTTG 180
DB 121 GGAATGGGACAGAGGATGAAAACGCGAATTTAGAAAGTGAAGAAATCTTCGTTG 180
OY 121 GGAATGGGACAGAGGATGAAAACGCGAATTTAGAAAGTGAAGAAATCTTCGTTG 180
DB 121 GGAATGGGACAGAGGATGAAAACGCGAATTTAGAAAGTGAAGAAATCTTCGTTG 180
OY 181 TAAACGAAAGCTTTTGGCTTCGACAAAGAAACGCGAATCTTCCTTAACATGAGG 240
DB 181 TAAACGAAAGCTTTTGGCTTCGACAAAGAAACGCGAATCTTCCTTAACATGAGG 240
OY 241 AGGAGAAAACATGAAAACAAAACCTGATTTTCTAATAATGAAGAGTCTGAAGACGAT 300
DB 241 AGGAGAAAACATGAAAACAAAACCTGATTTTCTAATAATGAAGAGTCTGAAGACGAT 300
OY 301 TGTCAATGTCGACCGCTTATGATTTATCCGACGCTAACTTGTGTAACAAGCGAGTTGA 360
DB 301 TGTCAATGTCGACCGCTTATGATTTATCCGACGCTAACTTGTGTAACAAGCGAGTTGA 360
OY 361 CATGATTTTAACTGCGTGAATTCCTTGAATGCTGCTCGGCTTGAATCTGCTGG 420
DB 361 CATGATTTTAACTGCGTGAATTCCTTGAATGCTGCTCGGCTTGAATCTGCTGG 420
OY 421 TGTGACAGTGGGACATGATCATATCAAAAAGCGGTTAAAGGGGTGCGCGAATAC 480
DB 421 TGTGACAGTGGGACATGATCATATCAAAAAGCGGTTAAAGGGGTGCGCGAATAC 480
OY 481 CTTTATTTGTGACAG 494
DB 481 CTTTATTTGTGACAG 494

RESULT 2
AAS02314
ID AAS02314 standard; DNA; 2363 BP.
XX AAS02314;
AC
XX 18-JUL-2001 (first entry)
DE
XX B. subtilis panBCD operon.
XX
KW PanBCD operon; PanC; PanB; pantothenate biosynthesis; vitamin B5;
KW nutritional supplement; panto-compound; pantoate; ds.
XX
OS Bacillus subtilis.

XX Key Location/Qualifiers
FH 242..1075
FT CDS /*cag= a
FT /*product= "PanB"
FT CDS 1077..1937
FT /*cag= b
FT /*product= "Panc"
FT CDS 1939..2332
FT /*cag= c
FT /*product= "Pand"

WO200121772-A2.
29-MAR-2001.
21-SEP-2000; 2000WO-US025993.
21-SEP-1999; 99US-00400494.
PR 07-JUN-2000; 2000US-0210072P.
PR 28-JUL-2000; 2000US-0221836P.
PR 24-AUG-2000; 2000US-0227860P.
XX (OMNI-) OMNIGENE BIOPRODUCTS.
PA
PI Yocum RR, Paterson TA, Hermann T, Pero JG;
XX WPI: 2001-218644/22.
DR P-PSDB; AAU01244, AAU01245, AAU01246.
XX
XX New recombinant microorganism which overexpresses a Bacillus subtilis
PT pantothenate biosynthetic enzyme, useful for the high yield production of
PT panto-compounds such as pantothenate and pantoate.
PS Claim 86; Page 198-201; 292pp; English.
XX
XX The sequence represents the B. subtilis PanBCD operon encoding PanB, PanC
CC and PanD, enzymes of the pantothenate biosynthetic pathway. Pantothenate,
CC also known as vitamin B5, is used as a nutritional supplement in mammals
CC and humans. The invention concerns methods of producing recombinant
CC microorganisms overexpressing at least one Bacillus subtilis pantothenate
CC biosynthetic enzyme. The microorganisms and methods of producing them are
CC useful for producing a panto-compound such as pantothenate or pantoate,
CC which is a nutritional requirement for livestock and humans. The methods
CC are also useful for the identification of pantothenate kinase modulators.
CC Panto-compounds are produced at a significantly higher yield than prior
CC art methods and can be produced independent of the need to feed
CC precursors which decreases expense
CC
SQ Sequence 2363 BP; 730 A; 451 C; 593 G; 589 T; 0 U; 0 Other;

Query Match 98.2%; Score 485; DB 4; Length 2363;
Best Local Similarity 100.0%; Pred. No. 4.6e-140;
Matches 485; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 TTGGTACAGCCCGTTGATTTGGTATCTTCATTTGGGACGATATCGCTGCAACTGCA 69
DB 1 TTGGTACAGCCCGTTGATTTGGTATCTTCATTTGGGACGATATCGCTGCAACTGCA 60
OY 70 CCTATTATTAATAATAGATGACATTCGACGAGTCTGCTTGAATCAAAAAGAGCTGGGA 129
DB 61 CCTATTATTAATAATAGATGACATTCGACGAGTCTGCTTGAATCAAAAAGAGCTGGGA 120
OY 130 CAGAGGATGAAATCGCCGCACTTTGAAAGTGAAGAAATCTTCGTTGAACGGAAG 189
DB 121 CAGAGGATGAAATCGCCGCACTTTGAAAGTGAAGAAATCTTCGTTGAACGGAAG 180
OY 190 GTTTTGGCTTGGCAGAGAAAACGCGACATCTCTCTTAACATGAGAGAGAGAAA 249
DB 181 GTTTTGGCTTGGCAGAGAAAACGCGACATCTCTCTTAACATGAGAGAGAGAAA 240
OY 250 CATGAAAACAAAACCTGATTTTCTAATAATGAAGAGTCTGAAGAACGATTTGTCATGCT 309

Db 241 CATGAAACAAACTGATTTCTTAAATGAGAGCTGAGAACCGATTGTCAGCT 300
 Qy 310 GACCGTTATGATTAATCCGAGCACTAAACCTTGAAACAAGCGAGTTGACATGATTT 369
 Db 301 GACCGTTATGATTAATCCGAGCACTAAACCTTGAAACAAGCGAGTTGACATGATTT 360
 Qy 370 AGTGGTGGATTCACCTTGGAAATGCTGCTCTGGCTTGATTCACCTGCTGGTGAAGT 429
 Db 361 AGTGGTGGATTCACCTTGGAAATGCTGCTCTGGCTTGATTCACCTGCTGGTGAAGT 420
 Qy 430 TGGGACATGATTCATCATCAAAAGCCGTTAAAGGGGCTGCGCAATACCTTTATTTG 489
 Db 421 TGGGACATGATTCATCATCAAAAGCCGTTAAAGGGGCTGCGCAATACCTTTATTTG 480
 Qy 490 GACAG 494
 Db 481 GACAG 485
 RESULT 3
 AAS02335/c
 ID AAS02335 standard; DNA; 7381 BP.
 XX
 AC AAS02335;
 XX
 DT 06-AUG-2003 (revised)
 DT 18-JUL-2001 (first entry)
 XX
 DE Plasmid pAN006 carrying the B. subtilis PanBCD operon.
 XX
 KW PanBCD operon; pantothenate biosynthesis; cyclic; circular; pAN006;
 KW vitamin B5; nutritional supplement; panto-compound; pantoate; P15; RBS1;
 KW ribosome binding site; ds.
 XX
 OS Bacteriophage.
 OS Escherichia coli.
 OS Bacillus subtilis.
 OS Chimeric.
 XX
 PN WO200121772-A2.
 XX
 PD 29-MAR-2001.
 XX
 PF 21-SEP-2000; 2000WO-US025993.
 XX
 PR 21-SEP-1999; 99US-00400494.
 PR 07-JUN-2000; 2000US-0210072P.
 PR 28-JUL-2000; 2000US-0221836P.
 PR 24-AUG-2000; 2000US-0227860P.
 XX
 PA (OMNI-) OMNIGENE BIOPRODUCTS.
 XX
 PI Yocum RR, Patterson TA, Hermann T, Pero JG;
 DR WPI, 2001-218644/22.
 XX
 PT New recombinant microorganism which overexpress a Bacillus subtilis
 PT pantothenate biosynthetic enzyme, useful for the high yield production of
 PT panto-compounds such as pantothenate and pantoate.
 XX
 PS Example 1; Page 287-292; 292pp; English.
 XX
 CC The sequence, pAN006, is a plasmid carrying the B. subtilis PanBCD operon
 CC which encodes three enzymes of the pantothenate biosynthetic pathway
 CC (PanB, C and D), under the control of the P15 promoter and using the
 CC artificial ribosome binding site, RBS1. Pantothenate, also known as
 CC vitamin B5, is used as a nutritional supplement in mammals and humans.
 CC The invention concerns methods of producing recombinant microorganisms
 CC overexpressing at least one B. subtilis pantothenate biosynthetic enzyme.
 CC The microorganisms and methods of producing them are useful for producing
 CC a panto-compound such as pantothenate or pantoate, which is a nutritional
 CC requirement for livestock and humans. The methods are also useful for the
 CC identification of pantothenate kinase modulators. Panto-compounds are

CC produced at a significantly higher yield than prior art methods and can
 CC be produced independent of the need to feed precursors which decreases
 CC expense. (Updated on 06-AUG-2003 to correct OS field.)
 CC
 XX Sequence 7381 BP; 1899 A; 1823 C; 1601 G; 2058 T; 0 U; 0 Other;
 SO
 Query Match 52.2%; Score 258; DB 4; Length 7381;
 Best Local Similarity 100.0%; Pred. No. 3.6e-69;
 Matches 258; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 237 GAGGAGGAG 296
 Db 3652 GAGGAGGAG 3593
 Qy 297 CGATTGTCATGCTGACCGCTTATGATTAATCCGAGCACTAACTTGCTGAACAAGCGGAG 356
 Db 3592 CGATTGTCATGCTGACCGCTTATGATTAATCCGAGCACTAACTTGCTGAACAAGCGGAG 3533
 Qy 357 TTGACATGATTTTATGCTGATTCACCTTGAATGCTGCTCTGGCTTGATTCACCTG 416
 Db 3532 TTGACATGATTTTATGCTGATTCACCTTGAATGCTGCTCTGGCTTGATTCACCTG 3473
 Qy 417 TCGGTGTCAGAGTTGCGGACATGATCCATCATACAAAAGCGTTAAAGGGGTCGCCGA 476
 Db 3472 TCGGTGTCAGAGTTGCGGACATGATCCATCATACAAAAGCGTTAAAGGGGTCGCCGA 3413
 Qy 477 ATACCTTATTTGTCACAG 494
 Db 3412 ATACCTTATTTGTCACAG 3395
 RESULT 4
 AAL48885
 ID AAL48885 standard; DNA; 6725 BP.
 XX
 AC AAL48885;
 XX
 DT 24-OCT-2002 (first entry)
 XX
 DE B subtilis PanB gene vector pAN636.
 XX
 KW 3-(2-hydroxy-3-methyl-butrylamino)-propionic acid; HMBPA; alpha-HIV;
 KW alpha-hydroxyisovalerate; pantothenate biosynthesis;
 KW cardiovascular disease; HMG CoA reductase; hypercholesterolemia;
 KW atherosclerosis; skin aging; PanB; PanD; PanE; PanF; vector; ds.
 XX
 OS Bacillus subtilis.
 OS Synthetic.
 XX
 PN WO200257476-A2.
 XX
 PD 25-JUL-2002.
 XX
 PF 19-JAN-2002; 2002WO-US001887.
 XX
 PR 19-JAN-2001; 2001US-0263035P.
 XX
 PA (OMNI-) OMNIGENE BIOPRODUCTS INC.
 XX
 PI Hermann T, Patterson TA, Pero JG, Yocum RR, Baldeusius K, Beck C;
 DR WPI, 2002-608383/65.
 XX
 PT Preparation of 3-(2-hydroxy-3-methyl-butrylamino)-propionic acid useful
 PT for synthesizing hydroxy methyl glutarate CoA reductase inhibitors,
 PT involves culturing a microorganism under suitable conditions and
 PT detecting or isolating product.
 XX
 PS Example 4; Page 72-74; 80pp; English.
 XX
 CC The present invention relates to a method of producing 3-(2-hydroxy-3-
 CC methyl-butrylamino)-propionic acid (HMBPA), which involves culturing a
 CC microorganism under conditions such that HMBPA is produced, and detecting

CC or isolating the HMBPA produced by the microorganism. These organisms may
CC have modified pantothenate biosynthesis enzymes. The method can be used
CC to produce HMBPA, which can then be used to synthesize inhibitors of
CC hydroxy methyl glutarate (HMG) CoA reductase, which are useful for
CC treating hypercholesterolemia, coronary atherosclerosis progression, and
CC to reduce risk of cardiovascular event in patients at risk. Alpha-HIV is
CC also produced, which is useful in the prevention of aging of skin, and to
CC treat skin disorders such as age spots, skin lines, wrinkles, photocaging
CC and aging. The present sequence is a vector used in the exemplification
CC of the invention

XX Sequence 6725 BP; 1773 A; 1541 C; 1668 G; 1743 T; 0 U; 0 Other;

XX Query Match 52.0%; Score 257; DB 6; Length 6725;

XX Best Local Similarity 100.0%; Pred. No. 7e-69; Mismatches 0; Indels 0; Gaps 0;

XX Matches 257; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 238 AGGAGGAGAAAAACATGAAAAAACAATGTTTCTTAAAAATGAAGAGTCTGAAGAAC 297
Db 325 AGGAGGAGAAAAACATGAAAAAACAATGTTTCTTAAAAATGAAGAGTCTGAAGAAC 384
Qy 298 GATTGTCAATGCTGACCGCTTATGATTATCCGAGAGCTAACTTGCTGAACAAGCGGAGT 357
Db 385 GATTGTCAATGCTGACCGCTTATGATTATCCGAGAGCTAACTTGCTGAACAAGCGGAGT 444
Qy 358 TGACATGATTTTAACTGCGTATTCACCTTGAATGCTGCTCGGCTTGAATCAACTGT 417
Db 445 TGACATGATTTTAACTGCGTATTCACCTTGAATGCTGCTCGGCTTGAATCAACTGT 504
Qy 418 CGGTGTGACAGTTGCGGACATGATCCATCATACAAAAGCCGTTAAAAAGGGGTGCGCGAA 477
Db 505 CGGTGTGACAGTTGCGGACATGATCCATCATACAAAAGCCGTTAAAAAGGGGTGCGCGAA 564
Qy 478 TACCTTTATTTGTGACAG 494
Db 565 TACCTTTATTTGTGACAG 581

RESULT 5

XX AAL41229
XX AAL41229 standard; DNA; 6725 BP.

XX AAL41229;

XX 30-OCT-2002 (first entry)

XX Pantothenate related vector PAN636 SEQ ID No 22.

XX 3-(2-hydroxy-3-methyl-butylamino)-propionic acid; HMBPA; PanB; PanB1;

XX pantothenate composition; deregulated pantothenate biosynthetic pathway;

XX deregulated isoleucine-valine; ilv biosynthetic pathway; PanB2; IlvC; ds.

XX unidentified.

XX WO200257474-A2.

XX 25-JUL-2002.

XX 19-JAN-2002; 2002WO-US001842.

XX 19-JAN-2001; 2001US-0262995P.

XX (OMNI-) OMNIGENE BIOPRODUCTS INC.

XX Hermann T, Patterson TA, Pero JG, Yocum RR, Baldenius K, Beck C;

XX WPI; 2002-636526/68.

XX Preparation of 3-(2-hydroxy-3-methyl-butylamino)-propionic acid-free
XX pantothenate composition by culturing microorganism having deregulated
XX pantothenate biosynthetic pathway and modified pantothenate biosynthetic
XX enzyme activities.

PS Example 4; Page 80-82; 93bp; English.

XX The invention relates to a method for the preparation of 3-(2-hydroxy-3-
XX methyl-butylamino)-propionic acid (HMBPA)-free pantothenate composition
CC by culturing microorganism having deregulated pantothenate biosynthetic
CC pathway and deregulated isoleucine-valine (ilv) biosynthetic pathway. The
CC microorganism has regulated activities of PanB; PanB1; PanB2; IlvC; PanB
CC and PanB1; PanB and PanB2; or, PanB and IlvC. The methods reduce HMBPA
CC (byproduct) production and provide increasing yields and purity of the
CC desired product. This polynucleotide sequence represents a vector
CC sequence relating to the pantothenate compositions of the invention

XX Sequence 6725 BP; 1773 A; 1541 C; 1668 G; 1743 T; 0 U; 0 Other;

XX Query Match 52.0%; Score 257; DB 6; Length 6725;

XX Best Local Similarity 100.0%; Pred. No. 7e-69; Mismatches 0; Indels 0; Gaps 0;

XX Matches 257; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 238 AGGAGGAGAAAAACATGAAAAAACAATGTTTCTTAAAAATGAAGAGTCTGAAGAAC 297
Db 325 AGGAGGAGAAAAACATGAAAAAACAATGTTTCTTAAAAATGAAGAGTCTGAAGAAC 384
Qy 298 GATTGTCAATGCTGACCGCTTATGATTATCCGAGAGCTAACTTGCTGAACAAGCGGAGT 357
Db 385 GATTGTCAATGCTGACCGCTTATGATTATCCGAGAGCTAACTTGCTGAACAAGCGGAGT 444
Qy 358 TGACATGATTTTAACTGCGTATTCACCTTGAATGCTGCTCGGCTTGAATCAACTGT 417
Db 445 TGACATGATTTTAACTGCGTATTCACCTTGAATGCTGCTCGGCTTGAATCAACTGT 504
Qy 418 CGGTGTGACAGTTGCGGACATGATCCATCATACAAAAGCCGTTAAAAAGGGGTGCGCGAA 477
Db 505 CGGTGTGACAGTTGCGGACATGATCCATCATACAAAAGCCGTTAAAAAGGGGTGCGCGAA 564
Qy 478 TACCTTTATTTGTGACAG 494
Db 565 TACCTTTATTTGTGACAG 581

RESULT 6

XX AAS02334/C
XX AAS02334 standard; DNA; 8503 BP.

XX AAS02334;

XX 06-AUG-2003 (revised)

XX 18-JUL-2001 (first entry)

XX Plasmid PAN004 carrying the B. subtilis PanBCD operon.

XX PanBCD operon; pantothenate biosynthesis; cyclic; circular; PAN004;

XX vitamin B5; nutritional supplement; panto-compound; pantoate; P26; RBS1;

XX ribosome binding site; ds.

XX Bacteriophage.

XX Escherichia coli.

XX Bacillus subtilis.

XX Chimeric.

XX WO200121772-A2.

XX 29-MAR-2001.

XX 21-SEP-2000; 2000WO-US025993.

XX 21-SEP-1999; 99US-00400494.

XX 07-JUN-2000; 2000US-021072P.

XX 28-JUL-2000; 2000US-0221836P.

XX 24-AUG-2000; 2000US-0227860P.

XX (OMNI-) OMNIGENE BIOPRODUCTS.

XX Yocum RR, Patterson TA, Hermann T, Pero JG;


```
XX OS Bacillus subtilis.
XX FH Key Location/Qualifiers
XX FT CDS 1..831
XX FT /*tag= a
XX FT /product= "ketopantoate hydroxymethyltransferase"
XX FT /partial
XX FT /note= "No stop codon"
XX
XX EN W0200121772-A2.
XX
XX PD 29-MAR-2001.
XX
XX PF 21-SEP-2000; 2000WO-US025993.
XX
XX PR 21-SEP-1999; 99US-00400494.
XX PR 07-JUN-2000; 2000US-0210072P.
XX PR 28-JUL-2000; 2000US-0221836P.
XX PR 24-AUG-2000; 2000US-0227860P.
XX
XX PA (OMNI-) OMNIGENE BIOPRODUCTS.
XX
XX PI Yocum RR, Patterson TA, Hermann T, Pero JG;
XX
XX DR WPI; 2001-218644/22.
XX DR P-PSDB; AAU01244.
XX
XX PT New recombinant microorganism which overexpress a Bacillus subtilis
XX PT pantothenate biosynthetic enzyme, useful for the high yield production of
XX PT panto-compounds such as pantothenate and pantoate.
XX
XX PS Claim 86; Page 169-170; 292pp; English.
XX
XX CC The sequence, panB, encodes B. subtilis ketopantoate
XX CC hydroxymethyltransferase, an enzyme of the pantothenate biosynthetic
XX CC pathway. Pantothenate, also known as vitamin B5, is used as a nutritional
XX CC supplement in mammals and humans. The invention concerns methods of
XX CC producing recombinant microorganisms overexpressing at least one Bacillus
XX CC subtilis pantothenate biosynthetic enzyme. The microorganisms and methods
XX CC of producing them are useful for producing a panto-compound such as
XX CC pantothenate or pantoate, which is a nutritional requirement for
XX CC livestock and humans. The methods are also useful for the identification
XX CC of pantothenate kinase modulators. Panto-compounds are produced at a
XX CC significantly higher yield than prior art methods and can be produced
XX CC independent of the need to feed precursors which decreases expense
XX
XX SQ Sequence 831 BP; 249 A; 166 C; 219 G; 197 T; 0 U; 0 Other;
XX
XX Query Match 49.4%; Score 244; DB 4; Length 831;
XX Best Local Similarity 100.0%; Pred. No. 3,1e-65;
XX Matches 244; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 251 ATGAAACAAACCTGATTTTCTAAATAAGAGAGTGAAGAACCGATTGTCATGCTG 310
XX DB 1 ATGAAACAAACCTGATTTTCTAAATAAGAGAGTGAAGAACCGATTGTCATGCTG 60
XX
XX QY 311 ACCGTTATGATTAATCCGGACGCTAAACTGTGTGAACAAGGGGAGTTGACATGATTTTA 370
XX DB 61 ACCGTTATGATTAATCCGGACGCTAAACTGTGTGAACAAGGGGAGTTGACATGATTTTA 120
XX
XX QY 371 GTCCGTGATTCACCTGGAATGTCCTCCGCGCTTGAATCAACTGTCGATGCACTT 430
XX DB 121 GTCCGTGATTCACCTGGAATGTCCTCCGCGCTTGAATCAACTGTCGATGCACTT 180
XX
XX QY 431 GCGGACATGATCCATCATACAAACCGTTAAAGGGGTGCGCCGCAATACCTTTATTTGTG 490
XX DB 181 GCGGACATGATCCATCATACAAACCGTTAAAGGGGTGCGCCGCAATACCTTTATTTGTG 240
XX
XX QY 491 ACAG 494
XX DB 241 ACAG 244
```

```
RESULT 9
ADN36367/C
ID ADN36367 standard; DNA; 10181 BP.
XX
XX AC ADN36367;
XX
XX DT 15-JUL-2004 (first entry)
XX
XX DE plasmid pDX17R for expression of the Bacillus subtilis yaadE sequence.
XX
XX KW B6 vitamin; Yaad; Yaaf; Epd; PdxA; PdxJ; PdxF; PdxB; PdxH; Dxs;
XX KW pyridoxine; pyridoxal; pyridoxamine; ds.
XX
XX OS Bacillus subtilis.
XX
XX EN W02004035010-A2.
XX
XX PD 29-APR-2004.
XX
XX PF 21-MAR-2003; 2003WO-US008880.
XX
XX PR 22-MAR-2002; 2002US-0367089P.
XX PR 25-MAR-2002; 2002US-0367863P.
XX PR 29-MAR-2002; 2002US-0368618P.
XX PR 03-MAR-2003; 2003US-00366618.
XX
XX PA (BMDI ) BASF AG.
XX
XX PI Yocum RR, Williams MK, Pero JG;
XX
XX DR WPI; 2004-365085/34.
XX
XX PT Producing B6 vitamin such as pyridoxine, pyridoxal or pyridoxamine,
XX PT involves culturing organism with increased yaad and/or yaaf activity as
XX PT compared to parent organism.
XX
XX PS Example 4; SEQ ID NO 8; 75pp; English.
XX
XX CC The invention relates to a method of producing (M1) a B6 vitamin, which
XX CC involves culturing an organism (I) with increased yaad and/or yaaf
XX CC activity, or Epd, PdxA, PdxJ, PdxF, PdxB, PdxH and/or Dxs activity as
XX CC compared to the parent organism, or a microorganism that overexpresses
XX CC Bacillus B6 vitamin biosynthetic gene, under conditions such that B6
XX CC vitamin is produced. (M1) is useful for producing B6 vitamin such as
XX CC pyridoxine, pyridoxal and pyridoxamine. This sequence represents the
XX CC plasmid pDX17R which contains the Bacillus subtilis yaadE operon used to
XX CC express these proteins.
XX
XX SQ Sequence 10181 BP; 2655 A; 2363 C; 2041 G; 3122 T; 0 U; 0 Other;
XX
XX Query Match 45.5%; Score 224.6; DB 12; Length 10181;
XX Best Local Similarity 92.5%; Pred. No. 1e-58;
XX Matches 236; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
XX
XX QY 1 GCTATATGTTGGTACAAAGCCGTTGATTTGGTATCACTTCATTGGGAGTATCGCTG 60
XX DB 3570 GCTATATGTTGGTACAAAGCCGTTGATTTGGTATCACTTCATTGGGAGTATCGCTG 3511
XX
XX QY 61 GAACTGCACCTATTTATTAATAATGATGACATTGACAGAGTGTGCTTGATCCAAAAA 120
XX DB 3510 GAACTGCACCTATTTATTAATAATGATGACATTGACAGAGTGTGCTTGATCCAAAAA 3451
XX
XX QY 121 GGAATGGGACAGAGGATGAAACTCCGCCGAACCTTTAGAAAGTGAAGAAATCCTTCGTTG 180
XX DB 3450 GGAATGGGACAGAGGATGAAACTCCGCCGAACCTTTAGAAAGTGAAGAAATCCTTCGTTG 3391
XX
XX QY 181 TAAAGGAAGGTTTTTGGCTTGGAGAAAGGAGATCATCTCCTTAACATGAGG 240
XX DB 3390 TAAAGGAAGGTTTTTGGCTTGGAGAAAGGAGATCATCTCCTTAACATGAGG 3331
XX
XX QY 241 AGGAAAAACATGAA 255
```

Dd		3330 GTTATAAGCATGAA 3316
	RESULT 10	
	AAL48883/c	
	ID AAL48883 standard; DNA; 6886 BP.	
XX		
AC	AAL48883;	
XX		
DT	24-OCT-2002 (first entry)	
XX		
DE	B subclis panBCD operon vector PAN624.	
XX		
OS	3-(2-hydroxy-3-methyl-butylamino)-propionic acid; HMBPA; alpha-HIV;	
XX	alpha-hydroxyisovalerate; pantothenate biosynthesis;	
KM	cardiovascular disease; HMg CoA reductase; hypercholesterolaemia;	
KW	atherosclerosis; skin aging; PanB; PanE; PanF; PanG; vector; ds.	
XX		
SS	Bacillus subtilis.	
OS	Synthetic.	
PN	WO200257476-A2.	
PD	25-JUL-2002.	
XX		
PF	19-JAN-2002; 2002MO-US001887.	
XX		
PR	19-JAN-2001; 2001US-0263053P.	
PA	(OMNI-) OMNIGENE BIOPRODUCTS INC.	
PI	Hermann T, Patterson TA, Pero JG, Yocum RR, Baldenius K, Beck C;	
DR	WPI; 2002-608383/65.	
XX		
PT	Preparation of 3-(2-hydroxy-3-methyl-butylamino)-propionic acid useful	
PT	for synthesizing hydroxy methyl glutarate CoA reductase inhibitors,	
PT	involves culturing a microorganism under suitable conditions and	
PS	detecting or isolating product.	
	Example 3; Page 68-70; 80pp; English.	
XX		
CC	The present invention relates to a method of producing 3-(2-hydroxy-3-	
CC	methyl-butylamino)-propionic acid (HMBPA), which involves culturing a	
CC	microorganism under conditions such that HMBPA is produced, and detecting	
CC	or isolating the HMBPA produced by the microorganism. These organisms may	
CC	have modified pantothenate biosynthesis enzymes. The method can be used	
CC	to produce HMBPA, which can then be used to synthesize inhibitors of	
CC	hydroxy methyl glutarate (HMg) CoA reductase, which are useful for	
CC	treating hypercholesterolaemia, coronary atherosclerosis progression, and	
CC	also produced, which is useful in the prevention of aging of skin, and to	
CC	treat skin disorders such as age spots, skin lines, wrinkles, photoaging	
CC	and aging. The present sequence is a vector used in the exemplification	
CC	of the invention	
SQ	Sequence 6886 BP; 1855 A; 1663 C; 1517 G; 1851 T; 0 U; 0 Other;	
	Query Match 43.8%; Score 216.2; DB 6; Length 6886;	
	Best Local Similarity 98.6%; Pred. No. 3.6e-56;	
	Matches 218; Conservative 0; Mismatches 3; Indels 0; Gaps 0;	
Oy	1 GGTATGTTGGTCAAGCCCGTTGATTGGTATCTTCATTGGGCAAGTAGTGCCG 60	
Dd	2909 GCTTAATGTTGGTATCAAGCCCGTTGATTGGTATCTTCATTGGGCAAGTAGTGCCG 2855	
Oy	61 CGAATCGCACCCTATTATTTAAATAGATAGACATTGCAGACGTCTGCCCTTGATCAA AAAA 120	
Dd	2849 CGAATCGCACCCTATTATTTAAATAGATAGACATTGCAGACGTCTGCCCTTGATCAA AAAA 2790	
Oy	121 GGACTGGGACAAGAGGATGAACCTGGCGCAACTTTAGAAGGAAAGAAATCCTTCGGTTG 180	
Dd	2789 GGAATGGGACAAGAGGATGAACCTGGCGCAACTTTAGAAGGAAAGAAATCCTTCGGTTG 2730	

QY	181	TAACGGAAGCTTTTGGCTTGCAGAGAAAGGCGAGATC	221
DB	2729	TAACGGAAGCTTTTGGCTTGCAGAGAAAGGCGAGATTC	2689
RESULT 11			
AL41227/c	AL41227 standard; DNA, 6886 BP.		
AC	AL41227;		
XX	30-OCT-2002	(first entry)	
DT	Pantothenate related vector pAN624 SEQ ID NO 20.		
XX	3-(2-hydroxy-3-methyl-butryl)amino)-propionic acid; HMBPA; Panb; Panb1;		
XX	pantothenate composition; deregulated pantothenate biosynthetic pathway;		
KM	deregulated isoleucine-valine; ilv biosynthetic pathway; Panb2; IlvC; ds-		
OS	Unidentified.		
XX	WO200257474-A2.		
PN	25-JUL-2002.		
XX	19-JAN-2002; 2002WO-US001842.		
PF	19-JAN-2001; 2001US-0262995P.		
XX	(OMNI-) OMNIGENE BIOPRODUCTS INC.		
PA	Hermann T, Patterson TA, Pero JG, Yocum RR, Baldeus K, Beck C;		
PI	WPI; 2002-636526/68.		
XX	Preparation of 3-(2-hydroxy-3-methyl-butryl)amino)-propionic acid-free		
XX	pantothenate composition by culturing microorganism having deregulated		
PT	pantothenate biosynthetic pathway and modified pantothenate biosynthetic		
PT	enzyme activities.		
XX	Example 3; Page 76-78; 93pp; English.		
PS	The invention relates to a method for the preparation of 3-(2-hydroxy-3-		
CC	methyl-butryl)amino)-propionic acid (HMBPA)-free pantothenate composition		
CC	by culturing microorganism having deregulated pantothenate biosynthetic		
CC	pathway and deregulated isoleucine-valine (ilv) biosynthetic pathway. The		
CC	microorganism has regulated activities of Panb1, Panb2; IlvC; Panb		
CC	and Panb1; Panb and Panb2; or, Panb and IlvC. The methods reduce HMBPA		
CC	(byproduct) production and provide increasing yields and purity of the		
CC	desired product. This polynucleotide sequence represents a vector		
CC	sequence relating to the pantothenate compositions of the invention		
XX	Sequence 6886 BP; 1855 A; 1663 C; 1517 G; 1851 T; 0 U; 0 Other;		
XX	Sequence 6886 BP; 1855 A; 1663 C; 1517 G; 1851 T; 0 U; 0 Other;		
QY	Query Match	43.8%; Score 216.2; DB 6; Length 6886;	
DB	Best Local Similarity	98.6%; Pred. No. 3.6e-56;	
DB	Matches 218; Conservative	0; Mismatches 3; Indels 0; Gaps 0;	
QY	1	GCTAATGTTGGTCAAGCCGTTGATTGGTATCTTCATTCAGGCAATGCGCTG	60
DB	2909	GCTAATGTTGGTCAAGCCGTTGATTGGTATCTTCATTCAGGCAATGCGCTG	2850
QY	61	CGAATCGCACTATTATTAATATAGATGACATTCGACAGTCTGCTTATCCAAAA	120
DB	2849	CGAATCGCACTATTATTAATATAGATGACATTCGACAGTCTGCTTATCCAAAA	2790
QY	121	GGAATCGGACAGAGGATGAATCTCGCGCACTTGAAGTGAAGATCCTTCTCGTTG	180
DB	2789	GGAATCGGACAGAGGATGAATCTCGCGCACTTGAAGTGAAGATCCTTCTCGTTG	2730
QY	181	TAACGGAAGCTTTTGGCTTGCAGAGAAAGGCGAGATC	221

Db 2729 IPACGGAAGTTTTTCGCTTCAGAGAAAACGGAAATTC 2699

RESULT 12
ID ABK73715 standard; DNA; 834 BP.
XX ABK73715
AC ABK73715;
XX
DT 13-AUG-2002 (first entry)
XX
DE Bacillus licheniformis genomic sequence tag (GST) #1006.
XX
KW Differential gene expression; genomic sequenced tag; GST;
XX altered culture condition; environmental stress;
XX physiological provocation; ds.
XX
OS Bacillus licheniformis.
XX
PN WO200229113-A2.
XX
PD 11-APR-2002.
XX
PF 05-OCT-2001; 2001WO-US031437.
XX
PR 06-OCT-2000; 2000US-00680598.
XX 27-MAR-2001; 2001US-0279526P.
XX
PA (NOVO) NOVOZYMES BIOTECH INC.
XX (NOVO) NOVOZYMES AS.
XX
PI Berka R, Clausen IG;
XX
DR WPI; 2002-416684/44.
XX

Monitoring differential expression of several genes in first Bacillus cell relative to expression of same genes in one or more second Bacillus cells, by using substrate containing Bacillus genomic sequenced tag array.

Claim 4; SEQ ID NO 1006; 200dp; English.

The invention describes a method of monitoring differential expression of genes in a first Bacillus cell relative to expression of the genes in other Bacillus cells, comprising hybridising labelled nucleic acid probes isolated from Bacillus cells to a substrate containing array of Bacillus genomic sequenced tags (GST), examining the array, and determining relative gene expression by an observed hybridisation reporter signal of a spot in the array. The method is useful for measuring the expression of a gene in a first Bacillus cell relative to expression of the same genes in one or more second Bacillus cells. The method is useful for monitoring global expression of several genes from a Bacillus cell, discovering new genes, identifying possible functions of unknown open reading frames and monitoring gene copy number variation and stability. Monitoring changes in expression of genes may be used to provide a representation of the way in which Bacillus cells adapt to changes in culture conditions, environmental stress or other physiological provocation. Extensive follow-up characterisation is unnecessary, when one spot on an array equals one gene or one open reading frame, since sequence information is available. This sequence represents a genomic sequence tag (GST) used in the method of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at http://wipo.int/pub/published_pct_sequences

Sequence 834 BP; 213 A; 186 C; 240 G; 195 T; 0 U; 0 Other;

Query Match 30.6%; Score 151.2; DB 6; Length 834;
Best Local Similarity 76.2%; Pred. No. 2,5e-36;
Matches 186; Conservative 0; Mismatches 58; Indels 0; Gaps 0

251 ATGAAAACAAACTGGATTTTCTTAAATAAGAGAGCTGAAAGAACCGATTCAGCTG 310
1 ATGAAGACAAATACTACTTTCTTAAATAAGAAACAGAGAGAGCGCATGCTCAAGCTG 60

Oy	311	ACCGCTTAAGTATATCCGCGACGTTAAATTCCTGTAAACAAGCGGAGATGTACATGATTTTA	370
Db	61	ACTGCTTAAGTATATCCGCGCGCAAAAGTTTCCCAACAGCGCGCGTGCATATGATCCTT	120
Oy	371	GTCGGGATTAACCTGGAATGGTCTGCTCCGCGCCTTGATTTCAACTGTCCGCTGTACAGTT	430
Db	121	GTCGGGATTAACCTGCGCATGGTGTGTATCCGCGCTTACTGCACCGATATCTGTTACGGTT	180
Oy	431	GCGGACATGATTCATCATACAAAGCCGTTAAAAGGGGTGGCCGAATACCTTATTGTG	490
Db	181	GCTGAATATGATCATCACCATACAAAGCTGTCAAAAGCGGCGCAAAAGATACGTTGTTGTG	240
Oy	491	ACAG 494	
Db	241	ACGG 244	
RESULT 13			
ID	ACA21619	standard; DNA; 837 BP.	
XX	ACA21619;		
AC	19-JUN-2003	(first entry)	
DT			
XX	Prokaryotic essential gene #3276.		
DE			
XX	Antisense; ds; prokaryotic essential gene; cell proliferation;		
KW	drug design; gene.		
RW			
XX	Bacillus anthracis.		
OS			
XX	W0200277183-A2.		
PN			
PD	03-OCT-2002.		
XX			
PF	21-MAR-2002; 2002WC-US009107.		
XX			
PR	21-MAR-2001; 2001US-00815242.		
PR	06-SEP-2001; 2001US-00948993.		
PR	25-OCT-2001; 2001US-0342923P.		
PR	08-FEB-2002; 2002US-00072851.		
PR	06-MAR-2002; 2002US-0362699P.		
XX			
PA	(EUIT-) ELITRA PHARM INC.		
XX			
P1	Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KU, Zyskind JW;		
P1	Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;		
XX			
DR	WPI: 2003-029926/02.		
DR	P-PSDB; ABU17749.		
XX			
PT	New antisense nucleic acid; useful for identifying proteins or screening		
PT	for homologous nucleic acids required for cellular proliferation to		
PT	isolate candidate molecules for rational drug discovery programs.		
XX			
PS	Claim 14; SEQ ID NO 9489; 1766bp; English.		
XX			
CC	The invention relates to an isolated nucleic acid comprising any one of		
CC	the 6213 antisense sequences given in the specification where expression		
CC	of the nucleic acid inhibits proliferation of a cell. Also included are:		
CC	(1) a vector comprising a promoter operably linked to the nucleic acid		
CC	encoding a polypeptide whose expression is inhibited by the antisense		
CC	nucleic acid; (2) a host cell containing the vector; (3) an isolated		
CC	polypeptide or its fragment whose expression is inhibited by the		
CC	antisense nucleic acid; (4) an antibody capable of specifically binding		
CC	the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular		
CC	proliferation or the activity of a gene in an operon required for		
CC	proliferation; (7) identifying a compound that influences the activity of		
CC	the gene product or that has an activity against a biological pathway		
CC	required for proliferation, or that inhibits cellular proliferation; (8)		
CC	identifying a gene required for cellular proliferation or the biological		

WP	AB069245_19	1900001	2010000
WP	AB069245_20	2000001	2110000
WP	AB069245_21	2100001	2210000
WP	AB069245_22	2200001	2310000
WP	AB069245_23	2300001	2410000
WP	AB069245_24	2400001	2510000
WP	AB069245_25	2500001	2610000
WP	AB069245_26	2600001	2710000
WP	AB069245_27	2700001	2810000
WP	AB069245_28	2800001	2910000
WP	AB069245_29	2900001	3010000
WP	AB069245_30	3000001	3011208

Query Match 27.7%; Score 137; DB 6; Length 110000;
Best Local Similarity 70.8%; Pred. No. 5.5e-31;
Matches 182; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY	238	AGAGAGAGAAAAATGAAAAACAAACCTGATTTTCTAAAAATGAAGAGCTGAAGAAC	297
Db	31566	AGAGCTAAGAAAAATGAAAAACAGTAGACTTTTGTATGAAGAAAAACGAGAAAA	31507
QY	298	GATTGTCAATGCTGACCGCTTATGATTATCCGGCAGCTAACTTGCTGAACAAGCGGGAGT	357
Db	31506	AATCAGATGATTACCGCTTACGATTATCTTCTGAAGATGTGAACAAGCAGAAAGC	31447
QY	358	TGACATGATTTTAAATCGGATTCCTGGAATGGTGGCTCGGCTTGATTCAACTGT	417
Db	31446	AGATATGATTTTAAATCGGATTCCTGGAATGGTGGCTTGATTCAACTGT	31387
QY	418	CGGTGTGACAGTTCGGAATGATTCATCAAAAGCCGTTAAAGGGGTGCGCCGAA	477
Db	31386	ACCGTAAATGATGATGATTCATCAATCAAAAGCCGTTAAAGGGGTGCGCCGAA	31327
QY	478	TACCTTTAATGTGACAG	494
Db	31326	TACTTTGTGTAAACGG	31310

Search completed: September 8, 2005, 20:33:22
Job time : 7325 secs

C	1	117.2	23.7	1880	4	US-09-710-279-4426	Sequence 4426, Ap
	2	111.6	22.6	825	3	US-09-334-001C-1569	Sequence 1569, Ap
	3	103.2	20.9	792	4	US-09-710-279-1417	Sequence 1417, Ap
C	4	99.6	20.2	4549	4	US-08-956-171E-1174	Sequence 174, App
	5	99.6	20.2	4549	4	US-08-781-986A-1174	Sequence 174, App
C	6	87.4	17.7	828	4	US-09-107-532A-359	Sequence 359, App
	7	87.4	17.7	828	4	US-09-334-000C-3123	Sequence 3123, App
C	8	78	15.8	3301	4	US-09-710-279-4209	Sequence 4209, Ap
	9	74.2	15.0	855	4	US-09-328-352-3820	Sequence 3820, Ap
C	10	61.2	12.4	9381	3	US-09-453-702B-7	Sequence 7, Appl
	11	59.2	12.0	912	4	US-09-252-991A-10351	Sequence 10351, A
C	12	59.2	12.0	1005	4	US-09-252-991A-10798	Sequence 10798, A
	13	54.2	11.0	889	4	US-09-902-540-6623	Sequence 6623, Ap
C	14	54.2	11.0	2364	4	US-09-902-540-506	Sequence 506, App
	15	54.2	11.0	640681	4	US-09-790-988-1	Sequence 1, Appl
C	16	52	10.5	2538	3	US-09-221-017B-369	Sequence 369, App
	17	50.2	10.3	845	4	US-09-543-681A-2061	Sequence 2061, Ap
C	18	49.2	10.0	813	4	US-09-489-039A-5090	Sequence 5090, Ap
	19	47.4	9.6	855	4	US-09-540-236-5	Sequence 5, Appl
C	20	47.4	9.6	2164	3	US-09-318-794A-1	Sequence 1, Appl
	21	47.4	9.6	2164	3	US-09-318-793A-3	Sequence 3, Appl
C	22	47.4	9.6	100848	4	US-09-596-002-39	Sequence 39, Appl
	23	45.8	9.3	4403765	3	US-09-103-840A-2	Sequence 2, Appl
C	24	45.8	9.3	4411529	3	US-08-998-416-570	Sequence 570, App
	25	43.8	8.9	753	3	US-09-949-019-1827	Sequence 1827, Ap
C	26	43.4	8.8	618	4	US-09-710-279-1827	Sequence 1827, Ap
	27	41.8	8.5	9454	4	US-09-949-016-12338	Sequence 12338, A

28	41.8	8.5	9455	4	US-09-949-016-14650	Sequence 14500, A
29	41.6	8.4	560	4	US-09-903-814A-1	Sequence 1, Appl1
30	41.6	8.4	560	4	US-10-723-061-1	Sequence 1, Appl1
C 31	39.2	7.8	8783	4	US-09-949-016-14402	Sequence 14402, A
C 32	38.6	7.8	7218	1	US-08-232-463-14	Sequence 14, Appl1
C 33	37.2	7.5	11652	4	US-09-818-512-3	Sequence 13, Appl1
34	36.8	7.4	601	4	US-09-949-016-172617	Sequence 172617, A
C 35	36.8	7.4	39259	4	US-09-949-016-16625	Sequence 16625, A
C 36	36	7.3	133559	4	US-09-949-016-15845	Sequence 15845, A
C 37	35.4	7.2	134808	4	US-09-949-016-156702	Sequence 156702, A
C 38	35.4	7.1	601	4	US-09-949-016-94870	Sequence 94870, A
C 39	35.2	7.1	30032	4	US-09-949-016-13933	Sequence 13933, A
C 40	35.2	7.1	38336	4	US-09-949-016-14502	Sequence 14502, A
C 41	35.2	7.1	58180	4	US-09-949-016-13383	Sequence 13383, A
C 42	35	7.1	10945	4	US-09-949-016-14097	Sequence 14097, A
C 43	34.8	7.0	601	4	US-09-949-016-15464	Sequence 15464, A
44	34.8	7.0	601	4	US-09-949-016-182713	Sequence 182713, A
45	34.8	7.0	67911	4	US-09-949-016-16579	Sequence 16579, A

ALIGNMENTS

```

RESULT 1
US-09-710-279-4426/c
; Sequence 4426, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUS4980US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4426
; LENGTH: 1880
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; US-09-710-279-4426

```

Query Match	23.7%	Score 117.2;	DB 4;	Length 1880;
Best Local Similarity	65.9%	Pred. No. 1.5e-27;		
Matches 170; Conservative	0;	Mismatches 88;	Indels 0;	Gaps 0

QY 232 AACATGAGAGGAGAAAACATGAAAACAAACTGGATTTCTTAAATAATGAGAGCTGA 291

Db 967 AATTGAGGGGTAAATGAAATTGAAAACCTTAAATCATTTTAAACAAATATGAGCATCACA 908

QY 292 AGAACCGATTGTCATGCTGACCGCTTAGTATATTCGCGACGCTAAACTTGCTGAACAAC 351

Db 907 GCAAAAGATTTCATGCGTTACAGCTTATGATTACTAGTGTCAACCAAGCAACAACAGC 848

QY 352 GGGAGTTGACATGATTTTAGTGGGATTCACGTGGAATGSGTGCTCGGCTTGATTC 411

Db 847 TGAAAATTGACATGATTTTGGTAGGAAATCTTTAGGAATGACAGTGTAGATATGATAG 768

QY 412 AACTGTCCGTGTGACAGTTCGCGACATGATTCATCAATACAAAAGCCGTTAAAGGGGTGC 471

Db 787 TACTGTTCAGTAACTTAATTAAACGATATGATTCATGTGAATGAGCTGTTPAAAAGAGGTGC 728

QY 472 GCCGAATACCTTTATGT 489

Db 727 TTCAGATCACTTTATAGT 710

RESULT 2
US-09-134-001C-1569

; Sequence 1569, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 1569
; LENGTH: 825
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1569

Query Match 22.6%; Score 111.6; DB 3; Length 825;
Best Local Similarity 66.8%; Pred. No. 6.4e-26;
Matches 159; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

Qy 252 TGAACAACTGATTTCTTAAATGAAGAGCTGAAGAACGATGTCATGCTGA 311
Db 8 TGAACCTTAAATCATTTAAACAAATGAAGAGCTACAGCAAAAGATTTCTATGCTTA 67
Qy 312 CCGCTTATGATTTATCCGAGCTAAATCTGCTGAACACGGGAGTTCATGATTTTAA 371
Db 68 CAGCTTATGATTTATCTAGTGTAAAGCAACCAAGCTAAATGACATGATTTTGG 127
Qy 372 TCGGATTTCACTTGAATGCTGCTCGGCTTGAATTCATGCTGCTGAGCACTTG 431
Db 128 TAGGAGATTTCTTGAAGATGACAGCTTGAAGATGATGATGATGATGATGATGATG 187
Qy 432 CGGACATGATTCATCATACAAAGCCGTTAAAGGGGCTGCCGGAATCCTTTATGT 489
Db 188 ACGATATGATTCATCATGATGTAAGGCTGTTAAAGAGGCTTCAATCATTTATAGT 245

RESULT 3
US-09-710-279-1417
; Sequence 1417, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P0348005
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1417
; LENGTH: 792
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-09-710-279-1417

Query Match 20.9%; Score 103.2; DB 4; Length 792;
Best Local Similarity 67.9%; Pred. No. 3.3e-23;
Matches 144; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

Qy 278 ATGAAGAGTCTGAAGAACCGATTCATGCTGACCGCTTATGATTCGGGAGCTAA 337
Db 1 ATGAAGGATCAGACCAAAAGTTTCTATGCTTAACGCTTATGATTCCTATGCTAAG 60
Qy 338 CTGCTGAACAACGGGAGTTGACATGATTTAGTGGTGAATTCATTGGAATGCTGTC 397

Db 61 CAAGCACAACAGCTGAATTTGACATGATTTGGTAGAGATTTCTTAGAATGACAGTG 120
Qy 398 CTGGCCCTGATTCATGCTGCGGTGACAGTTGCGGACATGATCCATCATACAAAGCC 457
Db 121 TTAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180
Qy 458 GTTAAAGAGGCTGCCGGAATACCTTTATGT 489
Db 181 GTTAAAGAGGCTGCTTCAGATACATTTATAGT 212

RESULT 4
US-08-956-171E-174/C
; Sequence 174, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; GIL H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Pannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 174:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4549 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 174:
US-08-956-171E-174

Query Match 20.2%; Score 99.6; DB 4; Length 4549;
Best Local Similarity 61.6%; Pred. No. 1.2e-21;
Matches 159; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

Qy 232 AACATGAGAGGAGAAAACATGAACAACTGATTTCTTAAATGAAGAGTCTGA 291
Db 4535 AAAAAGAGTTGATTAATCATTTGAACAGTAAATGATTAATGATGAACAAAGCA 4476
Qy 292 AGAACCATTTCTGCTGACCGCTTATGATTTATCCGAGCTTAAATTTGCTGAACAGC 351
Db 4475 AACTAAATTTCTATGATTAACGCTTATGATTTTCCAAAGTCTAAACAAAGTGAAAGCAGC 4416

Qy	352	GGAGGTGACATGATTTTAACTGCGGATTCACCTGGAAATGGCGCTCGGCGTGAATTC	411
Db	4415	GGGATTTGATATGATCTTGTGTGGGATTCCTGGTATGACTGATTAATGGTATGAAG	4356
Qy	412	AACGTGCGGTGACAGTTGCGGACATGATCCATCATCAAAAGCCGCTTAAAGGGGTGC	471
Db	4355	TACGGTACAAGACGTTAGCAGATATATTCATCAGGTGTCAGTAAAGAGGCCG	4296
Qy	472	GCCGAATACCTTATTTGT	489
Db	4295	ACCAATTCATTTGTAGT	4278

```

US-08-781-986A-174/c
: Sequence 174, Application US/08781986A
: Patent No. 6737248
: GENERAL INFORMATION:
: APPLICANT: Charles Kunisch
: TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
: NUMBER OF SEQUENCES: 5255
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Human Genome Sciences, Inc.
: STREET: 9410 Key West Avenue
: CITY: Rockville
: STATE: Maryland
: COUNTRY: USA
: ZIP: 20850
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
: COMPUTER: HP Vectra 486/33
: OPERATING SYSTEM: MSDOS version 6.2
: SOFTWARE: ASCII Text
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/781,986A
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Benson, Bob
: REGISTRATION NUMBER: 30,446
: REFERENCE/DOCKET NUMBER: PB248PP
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (301) 309-8504
: TELEFAX: (301) 309-8512
: INFORMATION FOR SEQ ID NO: 174:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 4549 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
US-08-781-986A-174

```

	Query Match	Similarity	20.2%;	Score 99.6;	DB 4;	Length 4549;
	Beet Local	Similarity	61.6%;	Pred. No. 1.2e-21;		
	Matches 159;	Conservative	0;	Mismatches 99;	Indels 0;	Gaps 0
Qy	232	AACATGAGGAGGAAAACATGAAAAAACAACCTGATTTCTTAAAAATGAGAGTCTGA	291			
Db	4535	AAAAAGAGGTGAATTAACATTGAAAAACATTAAGTCATTAATATGATATGAAAAAAGCA	447			
Qy	292	AGAACCGATTGTCATGCTCAACGGCTATATGATATATCCGAGAGTAACTTGCGAACAAC	351			
Db	4475	AACATAAATTTCTATGATGAACAGCTTATATATTTTCCAAGTGTAAACAAGTGAACAC	441			
Qy	352	GGAGAGTGCATGATTTTAGTCGGTGATTCACCTGGAAATGTCCTCGGCCCTTGATTC	411			
Db	4415	GGGATTTGATATGATTTCTTGTTGGGATTCACCTGGTATGACGTATTTAGTTATGAAG	435			
Qy	412	AACTGTCGGTGTGACAGTTGCGGACATGATTCATCATACAAAAGCCGTTAAAAAGGGTGC	471			

Db 4355 TACGGTCAAGATGACGTTAGCAGATGATTCATCAGCGTGTGCACTAGAGAGAGGCGC 4296
 Oy 472 GCCGATATCCTTTATGT 489
 Db 4295 ACCAAATGATTTGTAGT 4278

```

1      RESULT 6
2      US-09-107-532A-359
3      / Sequence 359, Application US/09107532A
4      / Patent No. 6583275
5      / GENERAL INFORMATION:
6      / APPLICANT: Lynn A Doucette-Stamm and David Bush
7      / TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
8      / ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
9      /
10     / NUMBER OF SEQUENCES: 7310
11     / CORRESPONDENCE ADDRESS:
12     / ADDRESSEE: GENOME THERAPEUTICS CORPORATION
13     / STREET: 100 Beaver Street
14     / CITY: Waltham
15     / STATE: Massachusetts
16     / COUNTRY: USA
17     / ZIP: 02354
18     / COMPUTER READABLE FORM:
19     / MEDIUM TYPE: CD-ROM ISO9660
20     / OPERATING SYSTEM: <Unknown>
21     / SOFTWARE: ASCII
22     / CURRENT APPLICATION DATA:
23     / APPLICATION NUMBER: US/09/107,532A
24     / FILING DATE: 30-Jun-1998
25     / PRIOR APPLICATION DATA:
26     / APPLICATION NUMBER: 60/085,598
27     / FILING DATE: 14 May 1998
28     / APPLICATION NUMBER: 60/051571
29     / FILING DATE: July 2, 1997
30     / ATTORNEY/AGENT INFORMATION:
31     / NAME: Arinello, Pamela Deneke
32     / REGISTRATION NUMBER: 40,489
33     / REFERENCE/DOCKET NUMBER: GTC-012
34     / TELECOMMUNICATION INFORMATION:
35     / TELEPHONE: (781)893-5007
36     / TELEFAX: (781)893-8277
37     / INFORMATION FOR SEQ ID NO: 359:
38     / SEQUENCE CHARACTERISTICS:
39     / LENGTH: 828 base pairs
40     / TYPE: nucleic acid
41     / STRANDEDNESS: double
42     / TOPOLOGY: circular
43     / MOLECULE TYPE: DNA (genomic)
44     / HYPOTHETICAL: NO
45     / ANTI-SENSE: NO
46     / ORIGINAL SOURCE:
47     / ORGANISM: Enterococcus faecium
48     / FEATURE:
49     / NAME/KEY: misc feature
50     / LOCATION: (B) LOCATION 1...828
51     / SEQUENCE DESCRIPTION: SEQ ID NO: 359:
52     /
53     / US-09-107-532A-359

```

	Query Match	Similarity	Score	DB #	Length
Best Local	Similarity	60.2%	Pred. No.	4,7e-18	
Matches	145	Conservative	0	Mismatches	96
				Indels	0
				Gaps	0
Qy	25	AAACAAACTGATTTTCTAAATATGAGGAGCTGAGAAACGATGTGATGCTGACC	313		
Db	7	AATACGCGGTACATTTTAAAGATCTTAAATTAGAAATGAAAAATTACATGCTTACA	66		
Qy	314	GCTTATGATTTATCCGCGAGCTAACTTGCTGACCAAGCGGAGTTGACATGATTTTATGCT	373		
Db	67	GCTTATGCTATTATCAACGCAAAATTAATTGATGTAAGCTGGGATTAATGGAATATTATAGTA	126		

GENERAL INFORMATION:
APPLICANT: Blactner, Frederick R.
Burland, Valerie
Perna, Nicole T.
Plunkett, Guy
Welch, Rod
TITLE OF INVENTION: No. 6365723e1 Sequences of E. coli O157
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53701-2113
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch. 1.44Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 8.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/453,702B
FILING DATE: 03-Dec-1999
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/110,955
FILING DATE: 04-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J.
REGISTRATION NUMBER: 27386
REFERENCE/DOCKET NUMBER: 960296.95017
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 251-5000
TELEFAX: (608) 251-9166
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 9381
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-453-702B-7

Query Match 12.4%; Score 61.2; DB 3; Length 9381;
Best Local Similarity 54.4%; Pred. No. 5.5e-09;
Matches 123; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 269 TTTCTAAATGAGAGAGCTTGAAGACCGATTGTCATGCTGACCGCTTATGATTATCCG 328
DB 1935 TTGCAAAATGTAAACGAAAAAAGCGCTCGCACCATCACCGCTACGACTACAGC 1876

QY 329 GCACTAAACTGCTGAACAGCGGAGTGAATGATTTAGTGGGATTCACCTTGA 388
DB 1875 TTCGAAAACTGTTCCGAGAGAGCGCTGAACGATGCTGCTGGCGACTGCTGGGG 1816

QY 389 ATGTCGTCCTCGGCTTGAATCACTGTCGCTGAGCAGTTGCGGACATGATCCAT 448
DB 1815 ATGACCGTTACAGGGGATTAATCTACCTTCCCGTTACCGTTAGAGATATGCCCTACAC 1756

QY 449 ACAAAGCCGTTAAAGGGGTCGCGCAATACCTTTATTTGACAG 494
DB 1755 ACTACCGAGTACGTCGTGGCGCGCAAACTGCTGCTGTTGGCTG 1710

RESULT 11
US-09-252-991A-10351
Sequence 10351, Application US/09252991A
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 10351
LENGTH: 912
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-10351

Query Match 12.0%; Score 59.2; DB 4; Length 912;
Best Local Similarity 54.6%; Pred. No. 7.2e-09;
Matches 118; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 279 TGAAGAGTCTGAAGAACCGATTGTCATGCTGACCGCTTATGATTATCCGACCTAAAC 338
DB 146 TGAAGCAGAGCGCGGAGAAAGTCCATGCTGACCTGACGACGACCTTGGCCACA 205

QY 339 TTGCTGAACAGCGGAGTGAATGATTTAGTGGTGAATTCATTGGAATGCTGTC 398
DB 206 CCGCAGCGAGCGCGGCTGAGCGTCTTGTGCGAAGATTCGCTCGAATGATATTC 265

QY 399 TCGGCTTGAATCACTGTCGCTGACAGTTGGGACATGATCATCAATCAAAAGCG 458
DB 266 AGGCTCAGACAGCACCCTGCGGTGACGAAAGAAATGCTATATCAACGCGCTGCG 325

QY 459 TTAAGGGGTCGCGCGCAATACCTTTATTTGACAG 494
DB 326 TGAACGGGTAAACAAAGTTCCTGATGTCACCG 361

RESULT 12
US-09-252-991A-10798/C
Sequence 10798, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 10798
LENGTH: 1005
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-10798

Query Match 12.0%; Score 59.2; DB 4; Length 1005;
Best Local Similarity 54.6%; Pred. No. 7.6e-09;
Matches 118; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 279 TGAAGAGTCTGAAGAACCGATTGTCATGCTGACCGCTTATGATTATCCGACGCTAAAC 338
DB 773 TGAAGCAGAGCGCGGAGAAAGTCCATGCTGACCTGACGACGACCTTGGCCACA 714

QY 339 TTGCTGAACAGCGGAGTGAATGATTTAGTGGTGAATTCATTGGAATGCTGTC 398
DB 713 CCGCAGCGAGCGCGGCTGAGCGTCTTGTGCGAAGATTCGCTCGAATGATATTC 654

QY 399 TCGGCTTGAATCACTGTCGCTGACAGTTGGGACATGATCATCAATCAAAAGCG 458
DB 653 AGGCTCAGACAGCACCCTGCGGTGACGAAAGAAATGCTATATCAACGCGCTGCG 594


```
Db 61 CGAAGTGCACCTATATTAATAATAGATGACATGTCAGACAGTCTGCTTGAATCCAAAAA 120
Qy 121 GGACCTGGACAGAGGAGATGAAATCTGCCGAATTTAGAAAAGTAAGATCTCTCCGTTG 180
Db 121 GGACTGGACAGAGGAGATGAAATCTGCCGAATTTAGAAAAGTAAGATCTCTCCGTTG 180
Qy 181 TAAAGGAGGTTTGGCTTGACAGAAAGAAAGGAGATCATCTCTCTAAGCATGAGG 240
Db 181 TAAAGGAGGTTTGGCTTGACAGAAAGAAAGGAGATCATCTCTCTAAGCATGAGG 240
Qy 241 AGGAGAAAACATGAAAACAAAACGATTTTCTAAAAATGAAGAGTCTGAAGAACGAT 300
Db 241 AGGAGAAAACATGAAAACAAAACGATTTTCTAAAAATGAAGAGTCTGAAGAACGAT 300
Qy 301 TGTCTGCTGACCGCTTATGATATCCGGCAGCTAACTTGCTGAACAAGGGAGTTGA 360
Db 301 TGTCTGCTGACCGCTTATGATATCCGGCAGCTAACTTGCTGAACAAGGGAGTTGA 360
Qy 361 CATGATTTTAACTGAGTTCCTGGAATGATGCTGCTCGGCTTGATTCAGTGTGCG 420
Db 361 CATGATTTTAACTGAGTTCCTGGAATGATGCTGCTCGGCTTGATTCAGTGTGCG 420
Qy 421 TGTGACAGTTGGCAGATGATCATATCAAAAACCGTTAAAAAGGGGTGCGCCGAAATAC 480
Db 421 TGTGACAGTTGGCAGATGATCATATCAAAAACCGTTAAAAAGGGGTGCGCCGAAATAC 480
Qy 481 CTTTATTTGTCAGAG 494
Db 481 CTTTATTTGTCAGAG 494
```

RESULT 2

```
US-10-984-449-59
; Sequence 59, Application US/10984449
; Publication No. US20050089973A1
; GENERAL INFORMATION:
; APPLICANT: Yocum, R. Rogers
; APPLICANT: Patterson, Thomas A.
; APPLICANT: Hermann, Theron
; APPLICANT: Petro, Janice G.
; TITLE OF INVENTION: METHODS AND MICROORGANISMS FOR PRODUCTION OF
; FILE REFERENCE: BGI-141PCN
; CURRENT APPLICATION NUMBER: US/10/984,449
; PRIOR FILING DATE: 2004-11-08
; PRIOR APPLICATION NUMBER: USSN 09/667,569
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: USSN 09/400,494
; PRIOR FILING DATE: 1999-09-21
; PRIOR APPLICATION NUMBER: USSN 60/210,072
; PRIOR FILING DATE: 2000-06-07
; PRIOR APPLICATION NUMBER: USSN 60/221,836
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: USSN 60/227,860
; PRIOR FILING DATE: 2000-08-24
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 59
; LENGTH: 2363
; TYPE: DNA
; ORGANISM: Bacillus subtilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (242)..(1072)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1077)..(1934)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1939)..(2319)
; US-10-984-449-59
```

```
Query Match 98.2%; Score 485; DB 21; Length 2363;
Best Local Similarity 100.0%; Pred. No. 5,8e-138;
Matches 485; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 10 TTGGTACAGAGCCGTTGATTTTGGTATATCTTCCATTTGGGCAAGTATCGCTGCGAATGCA 69
Db 1 TTGGTACAGAGCCGTTGATTTTGGTATATCTTCCATTTGGGCAAGTATCGCTGCGAATGCA 60
Qy 70 CCTATTTAAAAATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 129
Db 61 CTTATTTAAAAATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
Qy 130 CAGAGGATGAAATCTGCCGAACTTTAGAAAAGTGAAGATCTTCTGTTAAAGGAG 189
Db 121 CAGAGGATGAAATCTGCCGAACTTTAGAAAAGTGAAGATCTTCTGTTAAAGGAG 180
Qy 190 GTTTTGGCTTGCAAGAAAGAAAGGAGATCATCTCTCTTAAACATGAGAGAGAGAAA 249
Db 181 GTTTTGGCTTGCAAGAAAGAAAGGAGATCATCTCTCTTAAACATGAGAGAGAGAAA 240
Qy 250 CATGAAAACAAAACATGATTTTCTAAAAATGAAGAGTCTGAAGAACGATTTGTCATGCT 309
Db 241 CATGAAAACAAAACATGATTTTCTAAAAATGAAGAGTCTGAAGAACGATTTGTCATGCT 300
Qy 310 GACCGCTTAATGATATTCGGCAGCTAACTTGTGAACAAGCGGAGTTGACATGATTTT 369
Db 301 GACCGCTTAATGATATTCGGCAGCTAACTTGTGAACAAGCGGAGTTGACATGATTTT 360
Qy 370 AGTCGTGATTCACCTTGAATGCTGCTGCTGCGCTTGATTCATCTGCTGAGACAGT 429
Db 361 AGTCGTGATTCACCTTGAATGCTGCTGCTGCGCTTGATTCATCTGCTGAGACAGT 420
Qy 430 TGGCAGATGATCATCATATCAAAAACCGTTAAAAAGGGGTGCGCCGAAATCTTATTTGT 489
Db 421 TGGCAGATGATCATCATATCAAAAACCGTTAAAAAGGGGTGCGCCGAAATCTTATTTGT 480
Qy 490 GACAG 494
Db 481 GACAG 485
```

RESULT 3

```
US-10-984-449-94/c
; Sequence 94, Application US/10984449
; Publication No. US20050089973A1
; GENERAL INFORMATION:
; APPLICANT: Yocum, R. Rogers
; APPLICANT: Patterson, Thomas A.
; APPLICANT: Hermann, Theron
; APPLICANT: Petro, Janice G.
; TITLE OF INVENTION: METHODS AND MICROORGANISMS FOR PRODUCTION OF
; FILE REFERENCE: BGI-141PCN
; CURRENT APPLICATION NUMBER: US/10/984,449
; PRIOR FILING DATE: 2004-11-08
; PRIOR APPLICATION NUMBER: USSN 09/667,569
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: USSN 09/400,494
; PRIOR FILING DATE: 1999-09-21
; PRIOR APPLICATION NUMBER: USSN 60/210,072
; PRIOR FILING DATE: 2000-06-07
; PRIOR APPLICATION NUMBER: USSN 60/221,836
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: USSN 60/227,860
; PRIOR FILING DATE: 2000-08-24
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 94
; LENGTH: 7381
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant
```


Db 121 GTCGATGATTCATTGGAATGCTGCTCGGCTTGAATTCAGTGTGATGACAGTT 180
Qy 431 GCGGACATGATTCATCATATACAAAGCCGTTTAAAGGGGTGCGGAAATACCTTTATGTG 490
Db 181 GCGGACATGATTCATCATATACAAAGCCGTTTAAAGGGGTGCGGAAATACCTTTATGTG 240
Qy 491 ACAG 494
Db 241 ACAG 244

RESULT 9
US-10-508-768a-8/c
; Sequence 8, Application US/10508768A
; Publication No. US20050164335A1
; GENERAL INFORMATION:
; APPLICANT: YOCUM, R. ROGERS
; APPLICANT: WILLIAMS, MARK K.
; APPLICANT: PERO, JANICE G.
; TITLE OF INVENTION: METHODS AND ORGANISMS FOR PRODUCTION OF B6 VITAMERS
; FILE REFERENCE: OGI-00205
; CURRENT APPLICATION NUMBER: US/10/508,768A
; PRIOR FILING DATE: 2004-09-22
; PRIOR APPLICATION NUMBER: PCT/US03/008880
; PRIOR FILING DATE: 2003-03-21
; PRIOR APPLICATION NUMBER: 60/451,824
; PRIOR FILING DATE: 2003-03-03
; PRIOR APPLICATION NUMBER: 60/368,618
; PRIOR FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: 60/367,863
; PRIOR FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER: 60/367,089
; PRIOR FILING DATE: 2002-03-22
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 8
; LENGTH: 10181
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: plasmid - pdx17R
US-10-508-768a-8

Query Match 45.5%; Score 224.6; DB 22; Length 10181;
Best Local Similarity 92.5%; Pred. No. 1.3e-57;
Matches 236; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
Qy 1 GCTAATGTTGGTACAGCCCGTGTGATTTGTATCTTCATTCGATGGGAGATGCGCTG 60
Db 3570 GCTAATGTTGGTACAGCCCGTGTGATTTGTATCTTCATTCGATGGGAGATGCGCTG 3511
Qy 61 CGAATGCACTATTATTAATAATAGATAGACATTCAGAGATGCTGCTTGCATCAAAAAA 120
Db 3510 CGAATGCACTATTATTAATAATAGATAGACATTCAGAGATGCTGCTTGCATCAAAAAA 3451
Qy 121 GCACTGGGACAGAGGAGATGAACTGCGCACTTTAGAAAGTGAAGAAATCTTCTGTTG 180
Db 3450 GCACTGGGACAGAGGAGATGAACTGCGCACTTTAGAAAGTGAAGAAATCTTCTGTTG 3391
Qy 181 TAACGAAAGTTTTTGGCTTCGAGAAAGGCGAGATCATCTCTCTTAACATGAGG 240
Db 3390 TAACGAAAGTTTTTGGCTTCGAGAAAGGCGAGATCATCTCTCTTAACATGAGG 3331
Qy 241 AGGAGAAAAATGAA 255
Db 3330 GTTATTAAGGCAATGAA 3316

RESULT 10
US-10-466-642-20/c
; Sequence 20, Application US/10466642
; Publication No. US20040048343A1
; GENERAL INFORMATION:

; APPLICANT: OmniGene BioProducts
; TITLE OF INVENTION: METHODS AND MICROORGANISMS FOR THE PRODUCTION OF 3-(2-
; FILE REFERENCE: BGI-146PC
; CURRENT APPLICATION NUMBER: US/10/466,642
; CURRENT FILING DATE: 2003-07-18
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 6886
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: vector
; FEATURE:
; OTHER INFORMATION: PAN624
US-10-466-642-20

Query Match 43.8%; Score 216.2; DB 18; Length 6886;
Best Local Similarity 98.6%; Pred. No. 4e-55;
Matches 218; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 GCTAATGTTGGTACAGCCCGTGTGATTTGTATCTTCATTCGATGGGAGATGCGCTG 60
Db 2909 GCTAATGTTGGTACAGCCCGTGTGATTTGTATCTTCATTCGATGGGAGATGCGCTG 2850
Qy 61 CGAATGCACTATTATTAATAATAGATAGACATTCAGAGATGCTGCTTGCATCAAAAAA 120
Db 2849 CGAATGCACTATTATTAATAATAGATAGACATTCAGAGATGCTGCTTGCATCAAAAAA 2790
Qy 121 GCACTGGGACAGAGGAGATGAACTGCGCACTTTAGAAAGTGAAGAAATCTTCTGTTG 180
Db 2789 GCACTGGGACAGAGGAGATGAACTGCGCACTTTAGAAAGTGAAGAAATCTTCTGTTG 2730
Qy 181 TAACGAAAGTTTTTGGCTTCGAGAAAGGCGAGATC 221
Db 2729 TAACGAAAGTTTTTGGCTTCGAGAAAGGCGAGATC 2689

RESULT 11
US-10-466-717-20/c
; Sequence 20, Application US/10466717
; Publication No. US20040086982A1
; GENERAL INFORMATION:
; APPLICANT: Omogene Bioproducts
; TITLE OF INVENTION: PROCESSES FOR ENHANCED PRODUCTION OF PANTOTHENATE
; FILE REFERENCE: BGI-148PC
; CURRENT APPLICATION NUMBER: US/10/466,717
; CURRENT FILING DATE: 2003-07-18
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 6886
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: vector
; FEATURE:
; OTHER INFORMATION: PAN624
US-10-466-717-20

Query Match 43.8%; Score 216.2; DB 18; Length 6886;
Best Local Similarity 98.6%; Pred. No. 4e-55;
Matches 218; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 GCTAATGTTGGTACAGCCCGTGTGATTTGTATCTTCATTCGATGGGAGATGCGCTG 60
Db 2909 GCTAATGTTGGTACAGCCCGTGTGATTTGTATCTTCATTCGATGGGAGATGCGCTG 2850
Qy 61 CGAATGCACTATTATTAATAATAGATAGACATTCAGAGATGCTGCTTGCATCAAAAAA 120
Db 2849 CGAATGCACTATTATTAATAATAGATAGACATTCAGAGATGCTGCTTGCATCAAAAAA 2790

Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
; SOFTWARE: PatentIn version 3.2
```

SEQ ID NO 121456
LENGTH: 6540
TYPE: DNA
ORGANISM: Control Sequence
US-10-956-157-121456

Query Match 41.1%; Score 203; DB 21; Length 6540;
Best Local Similarity 100.0%; Pred. No. 4.5e-51;
Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GCTATGTTGGTACCAAGCCGTTGATTGGTATTAATCTTCATTGGGCAAGTATGCGCTG	60
DB	6338	GCTAATGTTGGTACCAAGCCGTTGATTGGTATTAATCTTCATTGGGCAAGTATGCGCTG	6397
QY	61	CGAATGCACTTATTATTAATAATAGATGACATTGACAGCTTGCCTTGATCCAAAAA	120
DB	6398	CGAATGCACTTATTATTAATAATAGATGACATTGACAGCTTGCCTTGATCCAAAAA	6457
QY	121	GGAATGGGACAGAGGATGAATCTGCCGAATTTAGAAAGTGAAGATCCTTCGCTG	180
DB	6458	GGAATGGGACAGAGGATGAATCTGCCGAATTTAGAAAGTGAAGATCCTTCGCTG	6517
QY	181	TACGGAAGGTTTTTGGCTTGC	203
DB	6518	TACGGAAGGTTTTTGGCTTGC	6540

Search completed: September 8, 2005, 23:56:05
Job time : 649 secs

This Page Blank (uspic)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 8, 2005, 16:54:28 ; Search time 3136 Seconds
(without alignments)
5996.093 Million cell updates/sec

Title: US-09-925-824A-1
Perfect score: 494
Sequence: 1 gccaatgctgtgtacacg.....gaacaccttattgtgcag 494

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_g881:*
9: gb_g882:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	79.2	16.0	772	2	BE641179 Crl2.2.10
2	70.8	14.3	700	7	CV032340 RTNACL1.7
3	69.8	14.1	553	4	BJ178509 BJ178509
4	68.6	13.9	525	4	BJ413963 BJ413963
5	66.4	13.4	709	7	CK447171 N8A5.SP6
6	61	12.3	550	7	CF507610 USDA-FP-1
7	61	12.3	599	7	CF506499 USDA-FP-1
8	60	12.1	684	6	CB687940 CBEST-22-B
9	59.2	12.0	1159	4	BZ567907 pac82-164
10	58.8	11.9	579	4	BZ567907 pac82-164
11	58.6	11.9	805	8	BJ029615 BJ029615
12	58.6	11.9	805	8	BJ029615 BJ029615
13	58.6	11.9	1252	3	CNS0A78Y CNS0A78Y
14	57.6	11.7	489	6	CA989682 EST643190
15	56.8	11.5	725	6	CD483545 atrol-26m
16	56.6	11.5	387	5	BQ493962 EST03128
17	56.6	11.5	447	5	BQ494730 EST03895
18	56.6	11.5	473	6	CA581877 EST001552
19	54.8	11.1	499	6	CB081675 hks5a09.5
20	54.8	11.1	585	8	BH398523 AG-ND-170
21	54.8	11.1	752	8	BH376769 AG-ND-134
22	54.2	11.0	703	5	BU065211 Pgr 6 P02
23	53.8	10.9	402	5	BU494310 Ljirnpent
24	53.6	10.9	1068	9	CNS06YTM AL421472 T3 end of

25	53	10.7	314	4	BG463405	BG463405 EM1_49 B0
26	53	10.7	481	4	BG411250	BG411250 EM1_27 E0
27	53	10.7	519	4	BG464232	BG464232 EM1_71 C0
28	53	10.7	528	2	BF176896	BF176896 EM1_4 A09
29	53	10.7	557	6	CD220973	CD220973 CCG1_72-B
30	53	10.7	656	6	CD224696	CD224696 CCG1_35-A
31	53	10.7	728	8	BH650859	BH650859 BOMCUB82TR
32	53	10.7	743	8	BZ448307	BZ448307 BOMCUB82TR
33	53	10.7	1200	9	CL958099	CL958099 BOMCUB82TR
34	52	10.5	885	7	CO006055	CO006055 BOMCUB82TR
35	51.8	10.5	578	1	AU235768	AU235768 AU235768
36	51.4	10.4	455	7	CF604653	CF604653 RADIC01_0
37	51.4	10.4	466	7	CF604658	CF604658 RADIC01_0
38	51.4	10.4	495	7	CF609965	CF609965 INF1001_0
39	51.4	10.4	661	8	AQ864621	AQ864621 n0b00023H
40	50.8	10.3	652	5	BQ471131	BQ471131 HV01D11T
41	50.8	10.3	653	1	AL504804	AL504804 AL504804
42	50.4	10.2	657	1	AL504805	AL504805 AL504805
43	50.2	10.2	416	8	BH410494	BH410494 1007018F0
44	50.2	10.2	447	4	BG275106	BG275106 WHE2117 H
45	50.2	10.2	538	7	CO525451	CO525451 3530_1_16

ALIGNMENTS

RESULT 1
LOCUS BE641179 772 bp mRNA linear EST 01-SEP-2000
DEFINITION Crl2.2.103 SP6 Ceratopteris Spore Library Ceratopteris richardii
LOCUS BE641179.1 GI:9958840
ACCESSION BE641179
VERSION BE641179.1
KEYWORDS EST.
SOURCE Ceratopteris richardii
ORGANISM Ceratopteris richardii
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Monilliformes; Filicophyta; Filicopsida; Filicales; Pteridaceae;
Ceratopteris.
1 (bases 1 to 772)
Chatterjee,A., San Miguel,P., Stout,S.C., Banks,J. and Roux,S.J.
Expressed sequence tags of cDNA clones from a C. richardii library
unpublished (2000)
COMMENT
Contact: Roux SJ
Section of Molecular Cell and Developmental Biology
University of Texas
Biology Building, Room 16, Austin, TX 78712, USA
Tel: 512 471 4238
Fax: 512 232 3402
Email: sroux@utcc.utexas.edu
Plate: Crl2.2 row: I column: 03
Seq primer: SP6.

FEATURES
source
1..772
Location/Qualifiers
/organism="Ceratopteris richardii"
/mol_type="mRNA"
/catalytic="Brogan"
/db_xref="taxon:49495"
/clone="Crl2.2.103"
/tissue_type="Gametophyte"
/cell_type="Spore"
/dev_stage="20 hours after germination initiation"
/clone_lib="Ceratopteris Spore Library"
/note="Vector: pCMVSPORT6; EST sequence from cDNA library.
cDNA library constructed from mRNA isolated from C.
richardii spores that had developed for 20 hours after
their germination had been initiated by white light."

ORIGIN
Query Match 16.0%; Score 79.2; DB 2; Length 772;
Best Local Similarity 62.8%; Pred. No. 7.9e-12;
Matches 123; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

Oy		294	AACGATGTGATGGTGA CCGCTTA TATGTTATTC CGAGC GTAAACTTG CTGTAACAAGCGG	353
Db		166	AACCATTACAATGCTA ACGGCCTTA TGACCTATCTT GCTGTGATG ATGTCATAGA GGCG	225
Oy		354	GAGTTGACATATTTT AGTCGTGAT TACTTAC TTGGAATGTCG TCCTCGGCTTGATCA A	413
Db		226	GCATTGACATCTGCTT GTCCGAGAT TCTTGCGGTA TGATGATGAT GCATATGATCA CA	285
Oy		414	CTGTGGTGTGACAGTT GCGGACATG ATCATCA AAAACCGTT AAAGGGGTGGC	473
Db		286	CATTACCTGTACAATG ATGATATGCTG CTCATTGTAGAGCAGTAGCA GAAGGTGCTA	345
Oy		474	CGAATACCTTTATGCT	489
Db		346	GACGATCTCTCTTGCT	361
RESULT 2				
CV032340				
LOCUS				
DEFINITION		CV032340	700 bp mRNA linear EST 23-AUG-2004	
ACCESSION		RTNACL1_7_A09.g2.A029	Roots plus added Nac1 Pinus taeda cDNA clone	
VERSION		CV032340	'5', mRNA sequence.	
KEYWORDS		CV032340.1	GI:51495163	
SOURCE		EST.		
ORGANISM		Pinus taeda (loblolly pine)		
TITLE		Pinus taeda		
JOURNAL		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
COMMENT		Spermatophytes; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.		
REFERENCE		1 (bases 1 to 700)		
AUTHORS		Pratt,L., Cordonnier-Pratt,M.-M., Lorenz,W.W., Zimmermann,C. and Dean,J.F.D.		
		An EST database from Nac1-created loblolly pine (Pinus taeda) roots unpublished (2004)		
		Other ESTs: RTNACL1_7_A09.b1.A029		
		Contact: Cordonnier-Pratt MM		
		Laboratory for Genomics and Bioinformatics		
		The University of Georgia, Department of Plant Biology		
		Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA		
		Tel: 706 542 1860		
		Fax: 706 583 0210		
		Email: mmp Pratt@uga.edu		
		RNA prepared and library constructed by W. Walter Lorenz (School of Forest Resources, University of Georgia); plant material prepared by Craig Zimmermann (School of Forest Resources, University of Georgia) using rooted cuttings provided by the Forest Biology Research Cooperative (FBRC) and the CLONES project at the University of Florida; sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.		
		Seq primer: JENREV (CAGGAACACTATGACC).		
FEATURES				
SOURCE				
		location/Qualifiers		
		1..700		
		/organism="Pinus taeda"		
		/mol_type="mRNA"		
		/strain="3 CLONES"		
		/db_xref="taxon:3152"		
		/clone="RTNACL1_7_A09.A029"		
		/lab_host="DH10B-T1 phage-resistant E. coli"		
		/clone_1lb="Roots plus added Nac1"		
		/note="Organ: Root; Vector: pSL180; Site 1: EcoRI; Site 2: XhoI; The library was prepared from polyA+ RNA from the roots of 1-year-old loblolly pine (Pinus taeda) cuttings that were rooted and then planted in washed sand. The rooted cuttings were maintained for 135 days (July 2003 harvest) under ambient conditions in a local greenhouse. They were kept on a weekly regimen of 0.5x nutrient-complete Hoagland's solution and supplemented with additional water sufficient to maintain a 15% soil moisture content. Twenty-four hours (24h) prior to harvesting roots for mRNA preparation, the potted trees		

Query Match	14.3%	Score 70.8	DB 7	Length 700
Best Local Similarity	55.8%	Pred. No. 2.2e-09		
Matches 135	Conservative 0	Mismatches 107	Indels 0	Gaps 0
248	AAACATGAAACAAACATCGATATTTCTTAAATGAAGAGAGCTGTGAAAGACGATGTGATG	307		
448	AAAGAGTCACTTAAACATCTTGCTCACTAAATATACAAAGGGAGCCCATTAACATG	507		
308	CTGACCGCTTATGATATTCGGGACGCTAAACCTTGCTGAACAAGCGGAGTTGACATGAT	367		
508	GTGATCGCTTATGATTTTCATCGGAGCTTCATGATGACAAATGCTGGCATTTGATATCTGC	567		
368	TTACTCGGTATTCATCTGAAATGAGTGTCTCGGCGCTTATTCATCTGTCGGTGACA	427		
568	TTACTAGGGAGTTAGTTGAAATGAGTGTGATGACATGACACAACTTACCTGTCACC	627		
428	GTTCGCGACATGATTCATACAAAGCCGTTTAAAGGGTGGCGCAATACCTTATT	487		
628	TTGATGATGATGCTGCTTCACTTGTGTGCTGTATCCCGTGGAGCATTTAATCTTCTT	687		
Qy	488 GT 489			
Db	688 GT 689			
RESULT 3				
Bu178509	553 bp	mRNA	linear	EST 16-OCT-2003
LOCUS	Bu178509	normalized full length cDNA library, chloronemata,		
DEFINITION	Bu178509	caldonemata and malformed buds Physcomitrella patens subsp. patens		
ACCESSION	Bu178509	cDNA clone pphb22k17 5', mRNA sequence.		
VERSION	Bu178509.1	GI:18346466		
KEYWORDS	EST.			
SOURCE	Physcomitrella patens subsp. patens			
ORGANISM	Physcomitrella patens subsp. patens			
REFERENCE	Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Bryophyta;			
AUTHORS	Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.			
	1 (bases 1 to 553)			
	Nishiyama,T., Fujita,T., Shin-I,T., Seki,M., Nishide,H.,			
	Uchiyama,I., Kamiya,A., Carninci,P., Hayashizaki,Y., Shinozaki,K.,			
	Kohara,Y. and Hasebe,M.			
	Comparative genomes of Physcomitrella patens gametophytic			
	transcriptome and Arabidopsis thaliana: implication for land plant			
	evolution			
	Proc. Natl. Acad. Sci. U.S.A. 100 (13), 8007-8012 (2003)			
TITLE				
JOURNAL				
MEDLINE				
PUBMED	12808149			
COMMENT	Contact: Tadabu Shin-i			
	Center For Genetic Resource Information			
	National Institute of Genetics			
	1111 Yatae, Mishima, Shizuoka 411-8540, Japan			
	Tel: 81-559-81-6856			
	Fax: 81-559-81-6855			
	Email: tehin@genes.nig.ac.jp			
	A backbone of the vector is basically from pBluescript II (KS),			
	that was in vivo excised from a 1-FPC phase vector (Carninci et al.,			
	2001). 5' end of the cDNA that was digested with XhoI was ligated			
	to SalI site of the vector and the 3' end including polyA tail was			
	ligated to BamH1 site of the			
	vector (5'- gagagagagagagatcacaaccttgagagatTTTTTTTTTTTIVN-3' was			
	used as a 1st 3' primer, and			
	5'-gggttcgagatcagctgttcacagacagcgatgactcgagaaacgannnnn-3' as 2nd			
	5'-hairpin primer, giving the following 5' boardest sequence,			
	AGGCCAAATCGCGGACGCTGCAATTCGTGAGAACCG). cDNA insert could be			
	amplified with conventional T7 and T3 primers. This full-length			
	cDNA library was generated according to the method described in			


```
/clone_11b="Aspergillus nidulans negative subtraction  
hybridization polysaccharide related cDNA plasmid library"  
/note="vector: pCMVSPORT6.0; Site 1: EcoRI; Site 2:  
HindIII, 5' end of cDNA cloned near EcoRI site of  
pCMVSPORT6.0 and 3' end cloned near HindIII site of  
pCMVSPORT6.0 Average length of insert is 1.49 kb"
```

Query Match	13.4%	Score 66.4;	DB 7;	Length 709;
Best Local Similarity	54.9%;	Pred. No. 4.2e-08;		
Matches 130; Conservative	0;	Mismatches 107;	Indels 0;	Gaps 0;

QY	247	AAAACGTAAACAAACCTGGAATTTCTTAAATAAGAGAGTCGTAAAGAACCGATTGCAT	306
Db	371	AAAGAAAGTTTACATGCAAGACCTTACGAATTTATACAAAGAAAGGCGAGCCTATTACAT	430
QY	307	GCTGACCGCTTATGATTTATCCGGCAGCTAAACTTCTGTAACAAGCGGAGATTGCATGAT	366
Db	431	GTTGACTGCGCAGATTTTCCGAGGCGCCCATGTCCCGACCGAGCGGAAATGGACATGAT	490
QY	367	TTTAACTCGGTGATTCACTTGGAAATGTCGTCTCGGACTTGAATTCACATGTCGATGAC	426
Db	491	TCTCGTGAGTATACCTTGGCAATGTCGCTCGGCATGACAGATACGAAGGAAATGAC	550
QY	427	AGTTGCGGACATGATTCATTCATACAAAGCCTTAAAGGGGTGGCCGAATACCTT	483
Db	551	TCTAATATACATGATAGTGACATGTCGATTTGTTTCGCGAGCTGTCTCAGAGGCGCTT	607

LOCUS	DEFINITION	LOCUS	DEFINITION
CE507610	550 bp mRNA	CE507610	550 bp mRNA
LOCUS	USA-FP_123000-111 Immature Ovaries from field-collected Valencia Sweet Orange (Citrus sinensis (L.) Osbeck) Citrus sinensis CDNA clone MVR-59, B02 5', mRNA sequence.	LOCUS	USA-FP_123000-111 Immature Ovaries from field-collected Valencia Sweet Orange (Citrus sinensis (L.) Osbeck) Citrus sinensis CDNA clone MVR-59, B02 5', mRNA sequence.

ACCESSION CFE07610
VERSION CFE07610.1 GI:34522794
KEYWORDS EST.
SOURCE Citrus sinensis
ORGANISM Citrus sinensis

REFERENCE
1 (bases 1 to 550)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; euroside II; Sapindales; Rutaceae; Citrus.

AUTHORS	TITLE
Chaparro, J. X., Bausher, M. G., Dang, P., Hunter, W. B., McKenzie, C. L., Niedz, R. P. and Shatters, R. G., Jr.	Expressed Sequence Tags from Immature Ovaries of Valencia Sweet Orange (<i>Citrus sinensis</i> (L.) Osbeck)

Unpublished (2003)
Contact: Chaparro, JX
Horticulture and Breeding Research Unit
USDA, ARS, U. S. Horticultural Research Laboratory
2001, South Rock Road, Fort Pierce, FL 34945, USA
Tel: 772 462 5830
Fax: 772 462 5986
Email: jchaparro@usbrl.ars.usda.gov
Seq primer: T3 Primer.

```

FEATURES
    source
        Location/Qualifiers
            1..550
                /organism="Citrus sinensis"
                /mol_type="mRNA"
                /cultivar="Valencia"
                /db_xref="taxon:2711"
                /clone="MYF-59_E02"
                /tissue_type="Immature ovaries"
                /clone_1fb="Immature Ovaries from field-collected Valencia Sweet Orange (Citrus sinensis (L.) Osbeck)"
                /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; Standard library construction protocols from Stratagene cDNA synthesis kit (cat.# 200401-5) and uni-ZAP XR vector kit (cat.# 237211) were followed using poly(A) RNA."

```

Query Match	12.3%	Score 61;	DB 7;	Length 550;
Best Local Similarity	56.1%	Pred. No. 1.5e-06;		
Matches 115;	Conservative 0;	Mismatches 90;	Indels 0;	Gaps 0;

QY	285	AGTCTGAGAA	CCGATTTGTC	GTCGTCAG	CCGCTTATG	TTATTCGG	CAGCTAA	CTTGCTG	344
Db	101	AGAACGAGAG	CCCATTTAC	CAATGTC	CCCGCTATG	CACTTCC	CTCGCGTG	CACCTCG	160
QY	345	AACAAGGGAG	AGTTGACAT	GATTTTATG	TGCGTATTC	ACTTGGAA	TGTTGTC	CTCGGCC	404
Db	161	ACAATGCGGG	CATTGATAT	TGTTTGTG	TCGATCTC	AGCGGC	CAATGTTG	TTTCACGTC	220
QY	405	TTGATTCAC	TGTCGCGT	GTGACAG	TTCGCGAC	TGATTCAT	CAATACAA	AGCCGTTAAA	464
Db	221	ACGCACTAT	CGTTGCCAT	CATCTCG	AGAGAAAT	GTTCAT	CTGCGCGG	CTGTGCTC	280
QY	465	GGGCTGGC	CGCAATAC	CTTTATTTG	T				
Db	281	GTGCGCTA	GAGACCG	TTACTTGT	305				

RESULT 7	CF506499	LOCUS	DEFINITION
	CF506499	599 bp	linear
		USDA-PP_121000-997	Immature Ovaries from field-collected Valencia
		Sheet Orange (Citrus sinensis (L.) Osbeck)	Citrus sinensis cDNA
		clone MPF-46 H05 5',	mRNA sequence.

ACCESSION	CF506499	GI:34521683
VERSION	CF506499.1	
KEYWORDS	EST.	
SOURCE	Citrus sinensis	
ORGANISM	Citrus sinensis	
	Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Sapindales; Rutaceae; Citrus.	

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1. (Pages 1, 10, 559)	Chaparro, J. X., Basher, M. G., Niedz, R. P. and Shatters, R. G. Jr.	Expressed Sequence Tags from Immature Ovaries of Valencia Sweet Orange (<i>Citrus sinensis</i> (L.) Osbeck)	Unpublished (2003)	Contact: Chaparro, JX

Horticulture and Breeding Research Unit
USDA, ARS, U. S. Horticultural Research Laboratory
2001, South Rock Road, Fort Pierce, FL 34945, USA
Tel: 772 462 5830
Fax: 772 462 5986
Email: jchaparro@ushr1.ars.usda.gov
Seq Primer: T3 Primer.

```

FEATURES
    source
        Location/Qualifiers
            1..599
                /organism="Citrus sinensis"
                /mol_type="mRNA"
                /cultivar="Valencia"
                /db_xref="taxon:2711"
                /clone="MVF-46 H05"
                /tissue_type="Immature ovaries"
                /clone_lib="Immature Ovaries from field-collected Valencia Sweet Orange (Citrus sinensis (L.) Osbeck)"
                /notes="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; Standard library construction protocols from Stratagene cDNA synthesis kit (cat.# 200401-5) and Uni-ZAP XR vector kit (cat.# 217211) were followed using poly(A) RNA."

```

ORIGIN	Query Match	Score 61	DB 7	Length 599
	Best Local Similarity	56.1%	Pred. No. 1,5e-06	
	Matches 115	Conservative	0	Mismatches 90
				Indels 0
				Gaps 0
QY	285	AGCTGAAGAACGATGTGATGCTGACGCGCTATGATTTATCCGGACGCTAAACTTGCTG	344	

Db 150 AGAAGGAGAGCCATTCATATGTCAACCGCTATGATATCCCTCGGCGGTGACCTCG 209

Qy 345 AACAAAGGAGAGTTGACATGATTTTACGTGGTATTCATTGAAATGCTGCTCGGCC 404

Db 210 AAGGTGGGCGATGATATATGTTTGGTGGTGAATCAGCGGCATGCTGCTTCAACGCTC 269

Qy 405 TTGATTCATCTGCTGGTGTGACATTCGCGACATGATTCATCAACAAAGCCGTTAAA 464

Db 270 ACGACATACGTTGTCCTGATCACTCTCGAGAAATGTTGTCATGTCGCGGCTGTGCTC 329

Qy 465 GGGGTGGCGCGAATACCTTTATGT 489

Db 330 GTGGCGCTAAGACACCTTACTTGT 354

RESULT 8
CB687940
LOCUS
DEFINITION CB687940 684 bp mRNA linear EST 04-SEP-2003
CBST-22-B-10 Mixed source, strain EP155 and EP155 infected with hypovirus CHV1-EP713 Cryphonectria parasitica cDNA clone EP155, EP155-CHV1-EP713 5-prime, mRNA sequence.

ACCESSION
VERSION CB687940 GI:34448809
KEYWORDS EST.
SOURCE Cryphonectria parasitica
ORGANISM Cryphonectria parasitica
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Sordariomycetidae; Diaporthales; Valsaceae; Cryphonectria-Endothia complex; Cryphonectria.

REFERENCE
AUTHORS 1 (bases 1 to 684)
Dawe,A.L., McMaine,V.C., Panglao,M., Kasahara,S., Chen,B. and Nuss,D.L.
An ordered collection of expressed sequences from Cryphonectria parasitica and evidence of genomic microsynteny with Neurospora crassa and Magnaporthe grisea
Microbiology 149 (9), 2373-2384 (2003)

JOURNAL
MEDLINE 22830414
PubMed 12949163

COMMENT
Contact: Dave AL, Nuss DL
Center for Biosystems Research
University of Maryland Biotechnology Institute
5115 Plant Sciences Building, College Park, MD 20742, USA
Tel: 301-405-7661 or 5111
Fax: 301-314-9075
Email: dave@umbi.umd.edu
Cryphonectria parasitica EST derived from mixed samples of cDNA from wild-type (strain EP155) and hypovirus-infected cultures.
Seq primer: M13-reverse

FEATURES
source
1. 684
Location/Qualifiers
/organism="Cryphonectria parasitica"
/mol_type="mRNA"
/strain="EP155"
/db_xref="ATCC (inhost):38755 and 52571"
/db_xref="taxon:5116"
/clone="EP155, EP155-CHV1-EP713"
/clone_1id="Mixed source, strain EP155 and EP155 infected with hypovirus CHV1-EP713"
/note="Cryphonectria parasitica EST derived from mixed samples of cDNA from wild-type (strain EP155) and hypovirus-infected cultures"

ORIGIN
Query Match 12.1%; Score 60; DB 6; Length 684;
Best Local Similarity 55.2%; Pred. No. 3.1e-06;
Matches 117; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

Qy 278 ATGAAGAGTCTGAAGAACGATGCTGACCGCTTATGATTAATCCGACCTAAA 337

Db 208 ATGTATCAGAAAGACGAGCCATCAACGATGACGAGCCATGACTCCGCTCCGCTGAC 267

Qy 338 CTTCGTGAACAAGCGGAGTTGATGATTTTATGTCGGTATTCCTTGATGATGCTGTC 397

Db 268 GTTGGCGATGACAGCGGCGATGAGATGATTCCTTGTGCTGAGACAGCTCGCATGTGCT 327

Qy 398 CTGGCTTATTCATCTGTGCTGTGAGAGTTGCGGACATGATTCATCAACAAAGCC 457

Db 328 CTGGGATGAGAGACAAAGTGAAGTGTGTGTGAGAGATGCTGTTGCACTGTCCAAGT 387

Qy 458 GTTAAAGGGGTGCGCGCAATACCTTTATGT 489

Db 388 GTTCCAGGCGGACAGAGTCCGATTTACCGT 419

RESULT 9
B2567907/c
LOCUS
DEFINITION B2567907 1159 bp DNA linear GSS 17-DEC-2002
pac82-164_7275.y2 pac82-164 Pseudomonas aeruginosa genomic clone
pac82-164_7275, genomic survey sequence.

ACCESSION
VERSION B2567907 GI:27200059
KEYWORDS GSS.
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas.

REFERENCE
AUTHORS 1 (bases 1 to 1159)
Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,B.E., Hastings,M., Burns,J.L., Kaul,R. and Olsen,M.V.
Whole-Genome-Sequence variation among multiple isolates of Pseudomonas aeruginosa library
J. Bacteriol. (2002) in press

JOURNAL
COMMENT Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymon@u.washington.edu
Class: shotgun.

FEATURES
source
1. 1159
Location/Qualifiers
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="2-164"
/db_xref="taxon:287"
/clone="pac82-164_7275"
/clone_1id="pac82-164"
/note="Clinical isolate 2-164 Whole genomic shotgun library."

ORIGIN
Query Match 12.0%; Score 59.2; DB 8; Length 1159;
Best Local Similarity 54.6%; Pred. No. 5.9e-06;
Matches 118; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

Qy 279 TGAAGAGTCTGAAGAACGATGCTGACCGCTTATGATTAATCCGACCTAAAAC 338

Db 350 TGAAGCAGAGCGGAGAAAGATCCGATGCTGATGAGAGCCACCTTCGCCACA 291

Qy 339 TTGCTGAACAAGCGGAGTTGACATGATTTTATGTCGGTATTCATTGGAATGCTGCTC 398

Db 290 CCGGACGCAAGCCGCGGTGACGCTGCTGTGTCGAGATTCCTCGGAATGATATGC 231

Qy 399 TCGGCTTATTCATCTGTGCTGTGACAGCTTGGCGGACATGATCATCATCAAAAGCG 458

Db 230 AGGCTACAGACGACCTGCGGTGAGCAACGAGAAATGCTTATCAACAGGCTGCG 171

Qy 459 TTAAGAGGGGTGCGCGCAATACCTTTATTTGAGAG 494

Db 170 TGAAGCGGCTAACAAAGTTCCTGATGCTGACCG 135

RESULT 10
B029615/c
LOCUS
B029615 579 bp mRNA linear EST 26-SEP-2003

DEFINITION B029615 NIBB Mochii normalized Xenopus neurula library Xenopus laevis cDNA clone X012p21 5', mRNA sequence.

ACCESSION B029615

VERSION B029615.1

KEYWORDS GI:17369985

SOURCE EST.

ORGANISM Xenopus laevis (African clawed frog)

Xenopus laevis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae; Xenopodinae; Xenopus; Xenopus.

1 (bases 1 to 579)

Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara,Y.

Expressed genes in X. laevis embryo

Unpublished (2001)

Contact: Tadasu Shin-i

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshin@genes.nig.ac.jp

The information of this clone is available through the following URL.

http://xenopus.nibb.ac.jp.

location/Qualifiers

1..579

/organism="Xenopus laevis"

/mol_type="mRNA"

/db_xref="taxon:8355"

/clone="X012p21"

/tissue_type="whole embryo"

/dev_stage="stage 15"

/clone_lib="NIBB Mochii normalized Xenopus neurula library"

ORIGIN

Query Match 11.9%; Score 58.8; DB 4; Length 579;

Best Local Similarity 54.7%; Pred. No. 6.6e-06;

Matches 117; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

QY 281 AAGAGTCTGAAGACCGATTGTCATGCTGACCGCTTATGATTATCCGACAGCTAACTT 340

DB 562 AAACGAGAAAAAAGCTTTCGCGACCATCACCGCTTATGACTATAGCTTGGCCAAATC 503

QY 341 GCTGAACAAGCGGAGTTGACATGATTTTATGTCGCTGATTCATCTTGAAATGCTCTCTC 400

DB 502 TTTGCTGATGAAGGAGCTTAACTCATGCTGCTGCGCGATTCGCTGGCATGACGCTTCAG 443

QY 401 GGCCTGATTTCAACTGCTGCTGCTGACGATTCGCGACATGATCATCAAAAGCGCTT 460

DB 442 GGGCAGCATCCACCTCGCAGCTAACCGTTGCCGATATGCTTACCACTGCGCCGCTA 383

QY 461 AAAAGGGGTGCGCCGATATCTTTTATGTCAG 494

DB 382 CGTCGCGCGACCAACTGCTGCTGCTGCTG 349

RESULT 11

BH236381/c 805 bp DNA linear GSS 13-NOV-2001

LOCUS BH236381

DEFINITION AUTGAS5TF AUTC Arabidopsis thaliana genomic clone AUTGAS5, genomic survey sequence.

ACCESSION BH236381

VERSION BH236381.1

KEYWORDS GI:16906739

SOURCE GSS.

ORGANISM Arabidopsis thaliana (thale cress)

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 805)

Tomu,C.D., Whitelaw,C.A., Pal,G., Van Aken,S.E., Uteback,T.V.,

TITLE Feldjlym,T.V. and Fraser,C.M. Survey sequencing of Arabidopsis thaliana BAC T1317

JOURNAL Unpublished (2001)

COMMENT Other_GSSs: AUTCA59TR

Contact: Chris Town

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA.

Tel: 301-838-3523

Fax: 301-838-0208

Email: cdtown@tigr.org

From Wash. U contig 1154.

Seq primer: TF

Class: sheared ends.

FEATURES

source

1..805

location/Qualifiers

1..805

/organism="Arabidopsis thaliana"

/mol_type="genomic DNA"

/ecotype="Columbia"

/db_xref="taxon:3702"

/clone="AUTCA58"

/clone_lib="AUTC"

/note="Vector: PHO82; Site 1: BstXI; 2-3 kb sheared BAC DNA inserted into PHO82 using BstXI linkers"

ORIGIN

Query Match 11.9%; Score 58.6; DB 8; Length 805;

Best Local Similarity 53.8%; Pred. No. 8.1e-06;

Matches 121; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

QY 265 GGAATTTCTAATAAGAGAGCTGTGAAGACCGATTGCTGACCGCTTATGATT 324

DB 380 GCAATTTAGACAGAGAGCTGTGAAGAGGAGCGATTACTATGATGACCCCTACGATT 321

QY 325 TCGCGACGTTAACTTCTGACACAGCGGAGTTGACATGATTTTGTGCGTATTCAC 384

DB 320 CCTTCGCGGTTCACATCATCATGCTGAATGAGTTGTCTCGCGAATTCAGC 261

QY 385 TGAATGCTGCTCGCTCGGCTTGAATTCATGCTGCTGTCAGATTCGCGACATGATCA 444

DB 260 TCGCATGTTGTTCATGATGATGACACACTCTTCCTATCTCTTGAAGAGATGCTGT 201

QY 445 TCATACAAAACCGTTAAAGGGGTGCGCGAATTCCTTTATGTT 489

DB 200 TCACTGTCGCGCGCTTCTCGCGAGCCAAAGACCACTTCTGT 156

RESULT 12

CNS0A78Y 1252 bp mRNA linear HTC 06-FEB-2004

LOCUS CNS0A78Y

DEFINITION Arabidopsis thaliana Full-length cDNA Complete sequence from clone GSUTP572FE06 of flowers and buds of strain col-0 of Arabidopsis thaliana (thale cress).

ACCESSION BX822672

VERSION BX822672.1

KEYWORDS GI:42466004

SOURCE HTC; GSIr cDNA.

ORGANISM Arabidopsis thaliana (thale cress)

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 1252)

Castelli,V., Aury,J.M., Jallion,O., Wincker,P., Clepet,C., Menard,M., Cruaud,C., Queier,F., Sarpelli,C., Schachter,V., Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.

Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation

Unpublished

2 (bases 1 to 1252)

Genoscope.

Direct Submission

Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :

BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

COMMENT

- Web : www.genoscope.cns.fr
 The sequences are based on single pass reads. members carried out Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.
 V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.
 URGV INRA : Clépet C., Caboche M.
 Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences) . 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full_length
<http://www.genoscope.cns.fr/cgi-bin/gdb/gdb?source=Arabidopsis.Location/Qualifiers>

FEATURES

source

1. .1252
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /strain="Col-0"
 /db_xref="taxon:3702"
 /clone="GSLTFB572F06"
 /tissue_type="Flowers and buds"
 /plasmid="PCWSP08T 6"
 complement(1. .1252)
 /gene="At3g61530"

gene

ORIGIN

Query Match 11.9%; Score 58.6; DB 3; Length 1252;
 Best Local Similarity 53.8%; Pred. No. 9e-06;
 Matches 121; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

Qy 265 GGATTTTCTAAATGAAGAGCTGAAGACCGATTGCTGACCGCTTATGATTA 324
 Db 131 GCAATTGAGACAGAGATCGAAAGGTGAGCCGATTAATGTCACCGCTACGATTA 190
 Qy 325 TCCGGACGCTAACTGCTGTAACAAGCGGAGTGAACATGATTTTAAGTCGATTA 384
 Db 191 CCTTCCTGCGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 250
 Qy 385 TGAATGCTGCTGCTGCGCTTGAATGATGATGATGATGATGATGATGATGATGAT 444
 Db 251 TGCATGTTTGTCCATGCTATGACACCACTTCTCTATCCTTAAGAGATGCTTGT 310
 Qy 445 TCATACAAAGCCGTTAAAGGGGTGCGCCGGAATACCTTTATTTGT 489
 Db 311 TCATGTCGCGCGCTTCTCGCGGAGCAAAAGACCACTTCTTGT 355

RESULT 13

CNS0A4X7 1678 bp mRNA linear HTC 06-FEB-2004
 LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone
 DEFINITION GSJLTS20207 of Adult vegetative tissue of strain col-0 of
 Arabidopsis thaliana (thale cress).
 ACCESSION BX823298.1 GI:42462911
 VERSION BX823298
 KEYWORDS HTC; GSJLT cDNA.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 1678)
 Castell V., Aury J.M., Jallion O., Wincker P., Clépet C.,
 Menard M., Cruaud C., Queller P., Scarpelli C., Schachter V.,
 Temple G., Caboche M., Weissenbach J. and Salanoubat M.
 Whole Genome Sequence Comparisons and Full-length cDNA Sequences:
 A Combined Approach to Evaluate and Improve Arabidopsis Genome
 Annotation
 Unpublished
 2 (bases 1 to 1678)
 JOURNAL Genoscope.
 AUTHORS Direct Submission
 TITLE

JOURNAL

COMMENT

Submitted (18-NOV-2003) Genoscope - Centre National de Séquençage :
 BP 191 91006 Evry cedex - FRANCE (E-mail : segref@genoscope.cns.fr)
 - Web : www.genoscope.cns.fr
 The sequences are based on single pass reads. members carried out Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.
 V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.
 URGV INRA : Clépet C., Caboche M.
 Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences) . 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full_length
<http://www.genoscope.cns.fr/cgi-bin/gdb/gdb?source=Arabidopsis.Location/Qualifiers>

FEATURES

source

1. .1678
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /strain="Col-0"
 /db_xref="taxon:3702"
 /clone="GSLTIS202D07"
 /tissue_type="Adult vegetative tissue"
 /plasmid="PCWSP08T 6"
 complement(1. .1678)
 /gene="At3g61530"

gene

ORIGIN

Query Match 11.9%; Score 58.6; DB 3; Length 1678;
 Best Local Similarity 53.8%; Pred. No. 9.7e-06;
 Matches 121; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

Qy 265 GGATTTTCTAAATGAAGAGCTGAAGACCGATTGCTGACCGCTTATGATTA 324
 Db 526 GCAATTGAGACAGAGATCGAAAGGTGAGCCGATTAATGTCACCGCTACGATTA 585
 Qy 325 TCCGGACGCTAACTGCTGTAACAAGCGGAGTGAACATGATTTTAAGTCGATTA 384
 Db 586 CCTTCCTGCGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 645
 Qy 385 TGAATGCTGCTGCTGCGCTTGAATGATGATGATGATGATGATGATGATGATGAT 444
 Db 646 TGCATGTTTGTCCATGCTATGACACCACTTCTCTATCCTTAAGAGATGCTTGT 705
 Qy 445 TCATACAAAGCCGTTAAAGGGGTGCGCCGGAATACCTTTATTTGT 489
 Db 706 TCATGTCGCGCGCTTCTCGCGGAGCAAAAGACCACTTCTTGT 750

RESULT 14

CA989682 489 bp mRNA linear EST 06-JAN-2003
 LOCUS EST643190 GLSD Medicago truncatula cDNA clone GLSD-42M4, mRNA
 DEFINITION sequence.
 ACCESSION CA989682.1 GI:27522576
 VERSION CA989682
 KEYWORDS EST.
 SOURCE Medicago truncatula (barrel medic)
 ORGANISM Medicago truncatula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
 Medicago.
 1 (bases 1 to 489)
 Grusak M.A., Sanac D., Town C.D., Van Aken S., Uterback T.,
 Cheung F. and Fraser C.M.
 ESTs from late stage developing seeds of Medicago truncatula
 Unpublished (2002)
 Contact: Grusak, M.A.
 USDA/ARS Children's Nutrition Research Center
 Baylor College of Medicine
 1100 Bates Street, Houston, TX 77030-2600, USA

Tel: 713 798 7044
 Fax: 713 798 7078
 Email: mgrusak@bcm.tmc.edu
 TIGR sequence name: MTRCF7ATX
 More information is available at: www.medicago.org
 Seq primer: SKmod (CTA GAA CTA GCG GAT CC).
 Location/Qualifiers

FEATURES
 source
 1..489
 /organism="Medicago truncatula"
 /mol_type="mRNA"
 /cultivar="Al7"
 /db_xref="taxon:3980"
 /clone="pGLSD-42M4"
 /tissue_type="Immature seeds"
 /dev_stage="25 to 35 days after pollination"
 /lab_host="XLOLR"
 /clone_1lb="GLSD"
 /note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; Immature seeds, collected from pods ranging in age from 25 to 35 days after pollination, were harvested from greenhouse-grown plants. Seed were removed and separated from pod walls and immediately frozen in liquid nitrogen. Seeds throughout the age range were pooled for mRNA extraction. cDNA was prepared from polyA+ enriched RNA. The cDNA was directionally ligated into the Unizap XR vector from Stratagene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-zap phage using Ex-assist helper phage and propagated in XLOLR cells."

ORIGIN

Query Match 11.7%; Score 57.6; DB 6; Length 489;
 Best Local Similarity 51.6%; Pred. No. 1.4e-05;
 Matches 132; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

224 CATGAGAGGAGAAACATGAAACAACTGATTTCTTAAATGAAGAGTCTGAAG 293
 118 CAACATCAAAAGAGATACACTTTCACACTGACAAACAAACATTAACCTCAC 177
 294 AACGATTTGATGCTGACCGCTTATGATATCCGCGACGTAACCTGCTGAACAGGG 353
 178 AGCCATACACATGCTCACTGCTTATGATACCTTCTGCTGTCACTCGATATGCGCG 237
 354 GAGTGAATGATTTTAACTGCGTGATTCATCTGAAATGCTGCTCGGCTTGAATCA 413
 238 CATTGATATCTGCTTGTGATGATTCCTCTCATGATGCTGATGATGATGATGAT 297
 414 CTGCGGTGTGACAGTTGCGGACATGATCCATCAACAAAGCCGTTAAAGGGGTGCGC 473
 298 CTTTCCATTAATCTTAAATGAAAGCTTGTTCATTTGCTGCTGCTGCTGCTGCTA 357
 474 CGAATACTTTATTTGT 489
 358 AAATCTCTCTCTCTGT 373

RESULT 15
 CD483545 725 bp mRNA linear EST 04-JUN-2003
 LOCUS atrol-26msl-d12 Atrol Amborella trichopoda cDNA clone
 DEFINITION atrol-26msl-d12 5', mRNA sequence.
 ACCESSION CD483545
 VERSION CD483545.1 GI:31404813
 KEYWORDS EST.
 SOURCE Amborella trichopoda
 ORGANISM Amborella trichopoda
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; stem Magnoliophyta; Amborellaceae; Amborella.
 1 (bases 1 to 725)
 dePamphilis,C., Soltis,D., Soltis,P., Farmerie,W., Ma,H., Tanksley,S., Leebens-Mack,J., Field,D., Buzgo,M., Kim,S., Ilut,D., Landherr,L., Hu,Y., Wall,K., Albert,V., Carlson,J., Doyle,J.,

Prohlich,M., Miller,W., Oppenheimer,D. and Theissen,G.
 Generation of ESTs from early flower buds of Amborella trichopoda
 Unpublished (2002)
 COMMENT
 TITLE
 JOURNAL
 CONTACT: Claude dePamphilis or James Leebens-Mack
 Mueller Laboratory
 Penn State University
 208 Mueller Laboratory, Department of Biology, ATTN Rm212, Penn
 State University, University Park, PA 16802, USA
 Tel: 814 865 6413
 Fax: 814 865 9131

Email: cwd3@psu.edu or jhl10@psu.edu
 The sequence provided is trimmed of vector and low quality regions.
 Full sequence and original trace file are available from the Plant
 Genome Network website (http://Pgn.cornell.edu)
 Plate: atrol-26msl row: d column: 12
 Seq primer: M13P.

FEATURES

source

Location/Qualifiers

1..725
 /organism="Amborella trichopoda"
 /mol_type="mRNA"
 /db_xref="taxon:13333"
 /clone="atrol-26msl-d12"
 /tissue_type="flower buds"
 /dev_stage="2.5mm buds"
 /lab_host="SOLR"
 /clone_1lb="Atrol1"
 /note="Vector: pBluescript SK (+/-); Site 1: EcoRI; Site 2: XhoI; Amborella trichopoda Ball; This library was made from male flowers only. Only floral buds with diameter of 2.5 mm or less were used for RNA isolation. This is a directionally cloned, non-normalized library. Avg. insert length: 1611; Primers: M13P and M13R; Antibiotic: 50 ug/ml Ampicillin; Primary Titer: 2.2486 pfu total; Amplified Titer: 1.37E10 pfu/ml; Mass Excised Titer: 3.53E9 total; This library has been generated by the Floral Genome Project (FGP). We would like to thank David Lorence at the National Tropical Botanical Garden for providing plant material for library building. The Floral Genome Project is funded by NSF's Plant Genome Research Program (DBI-0115684). More information about the project can be obtained at http://fgp.bio.psu.edu"

ORIGIN

Query Match 11.5%; Score 56.8; DB 6; Length 725;
 Best Local Similarity 60.3%; Pred. No. 2.7e-05;
 Matches 94; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

256 AACAAATCTGATTTCTTAAATGAAGAGTGAAGAACGATTTGCTGACCGC 315
 570 AACCTAAGAAATTTGGGCTGAAGCATTAAGAGGGGAGACCTATCACATGCTAC 629
 316 TTATGATTTATCCGCGACTAACTTGCTGAACAAGCGGAGTGACATGATTTAGTCG 375
 630 TTATGATTTATCCGCGACTGCTTCACTCGATCAAGCGGAATGATTTAGACTGTGG 689
 376 TGATTTCACTGGAATGCTGCTCTCGGCTTGATTC 411
 690 TGATTTGCAAGCATGCTGCTTCAATGGGATGATAC 725

Search completed: September 8, 2005, 22:05:28
 Job time : 3143 secs